CS170-1

Medical Care

March 1982
Vol. 20 No. 3
MDLCBD · ISSN 0025-7079

OFFICIAL JOURNAL OF THE MEDICAL CARE SECTION, AMERICAN PUBLIC HEALTH ASSOCIATION

Evaluation Methodology for Ambulatory Care Information Systems

How Good Are the Data? Reliability of One Health Care Data Bank

The Interaction Among the Criteria Physicians Use When Prescribing

Does Cost Information Availability Reduce Physician Test Usage? A Randomized Clinical Trial With Unexpected Findings

The Effects of Price Information on Physicians' Test-Ordering Behavior: Ordering of Diagnostic Tests

Physician Sensitivity to the Price of Diagnostic Tests: A U.S.–Canadian Analysis

Why Veterans Choose Veterans Administration Hospitalization: A Multivariate Model

Satisfaction With Maternity Care: A Matter of Communication and Choice

Utilization of Specialized Ambulatory Care by the Elderly: A Study of a Clinic

PUBLISHED BY

J. B. Lippincott Company · Philadelphia · Toronto

March 1982
Volume 20, Number 3

CONTENTS

POLICY FORUM

Medically Oriented Computer-Based Information Systems. *Alan F. Dowling, Jr.* 253

ORIGINAL ARTICLES

Evaluation Methodology for Ambulatory Care Information Systems. *Donald W. Simborg and Quinn E. Whiting-O'Keefe* 255

How Good are the Data? Reliability of One Health Care Data Bank. *Leslie L. Roos, Jr., Noralou P. Roos, Sandra M. Cageorge and J. Patrick Nicol* 266

The Interaction Among the Criteria Physicians Use When Prescribing. *Robert N. Zelnio* 277

Does Cost Information Availability Reduce Physician Test Usage? A Randomized Clinical Trial With Unexpected Findings. *David I. Cohen, Paul Jones, Benjamin Littenberg and Duncan Neuhauser* 286

The Effects of Price Information on Physicians' Test-Ordering Behavior: Ordering of Diagnostic Tests. *K. Michael Cummings, Kenneth B. Frisof, Michael J. Long and George Hrynkiewich* 293

Physician Sensitivity to the Price of Diagnostic Tests: A U.S.—Canadian Analysis. *John Hoey, John M. Eisenberg, Walter O. Spitzer and Duncan Thomas* 302

Why Veterans Choose Veterans Administration Hospitalization: A Multivariate Model. *William Frank Page* 308

Satisfaction With Maternity Care: A Matter of Communication and Choice. *Deborah A. Sullivan and Ruth Beeman* 321

Utilization of Specialized Ambulatory Care by the Elderly: A Study of a Clinic. *Stanley R. Ingman and Ian R. Lawson* 331

BOOK REVIEWS

Medical Records in Health Information, by Kathleen A. Waters and Gretchen Frederick Murphy. *Faye Brown* 339

Managing Health Records: Administrative Principles, by Joan Gratto Liebler. *Faye Brown* 339

Computers in Nursing, edited by Rita D. Zielstorff. *Barbara J. Horn* 340

Nursing Information Systems, edited by Harriet H. Werley and Margaret R. Grier. *Barbara J. Horn* 340

Improving Doctor Performance: A Study in the Use of Information and Organizational Change, by Steven Kelman. *Robert H. Barnes* 341

Ethnicity and Medical Care, edited by Alan Harwood. *Renate Wilson and Max Wilson* 343

Communicating Effectively, The Management Anthology Series, Theme Two: Management Functions, edited by Ann J. Huntsman and Jane L. Binger. *Ruth Lyn Riedel* 344

Health Care Management and Contemporary Perspective, by Seth B. Goldsmith. *John R. Griffith* 345

EDITORIAL BOARD

Donald C. Riedel, Ph.D.
Editor-in-Chief
University of Washington

Jan Van Dusen
Editorial Assistant

S. E. Berki, M.A.
University of Michigan

Robert H. Brook, M.D., Sc.D.
University of California
Los Angeles

Gordon H. De Friese, Ph.D.
University of North Carolina

Eugene Feingold, Ph.D.
University of Michigan

Brant E. Fries, Ph.D.
Yale University

Thomas R. Godkins, B.S., F.A.
University of Oklahoma

Harold Hunter, Dr. P.H.
Washington, D.C.

Sherman A. James, Ph.D.
University of North Carolina

Anthony L. Komaroff, M.D.
Brigham and Women's
Hospital, Boston

James P. LoGerfo, M.D., M.P.H.
University of Washington

Nancy Milio, Ph.D.
University of North Carolina

Duncan Neuhauser, Ph.D.
Case Western Reserve University

John E. Ware, Jr., Ph.D.
The Rand Corporation

Alonzo S. Yerby, M.D.
Harvard University

Alfred Yankauer, M.D.
University of Massachusetts
Medical School

MEDICAL CARE (ISSN 0025-7079) is published monthly by J. B. Lippincott Company, East Washington Square, Philadelphia, PA 19105. Subscription address: Journal Fulfillment Department, Lippincott/Harper, 2350 Virginia Avenue, Hagerstown, MD 21740. Subscription rates: United States and U.S. Possessions, one year for individuals, $45.00, for institutions, $50.00, special price U.S. students, $27.00; all other countries except Japan, one year, for individuals, $55.00, for institutions, $60.00. Single copies $6.00. Subscription orders and correspondence concerning subscriptions in Japan should be sent to Woodbell Scope, 11-11 Shoto 2-chome, Shibuya-ku, Tokyo 150, Japan. Printed in U.S.A. Copyright © 1982 by J. B. Lippincott Company. Second class postage paid at Hagerstown, Maryland, and additional mailing offices.

MEDICAL CARE
March 1982, Vol. XX, No. 3

How Good Are the Data?

Reliability of One Health Care Data Bank

LESLIE L. ROOS, JR., PH.D., NORALOU P. ROOS, PH.D.,
SANDRA M. CAGEORGE, S.M., AND J. PATRICK NICOL, B. COMM.

This study investigates the reliability of the Manitoba Health Services Commission data bank from a variety of perspectives. Emphasizing diagnostic and surgical procedures, the research focuses on those areas in which problems exist and in which the data can be relied upon. Computerized comparisons are stressed, since they can provide cost-effective checks on data quality. One key to performing reliability studies inexpensively is finding information recorded independently: by separate individuals or organizations, at two different times, or in two or more data files. When a particular event has certain logical implications vis-à-vis another, inconsistencies can be located. Face sheet information and data on the performance of major surgical procedures were found to be reliably recorded in the Manitoba data bank. Collapsing ICD-8 diagnoses from medical claims into several categories proved much better than relying upon individual diagnoses. Problems in working with the data included difficulty in distinguishing between closely related surgical procedures and the underreporting of inhospital consultations and nonsurgical procedures.

A MAJOR ISSUE with health care data banks concerns their reliability and validity for a wide spectrum of possible research questions. Other investigators have shown the suitability of Medicare/Medicaid data for the study of ambulatory quality of care,[1] of antibiotic usage[2] and of elective surgery.[3] On the other hand, Demlo et al.[4] have cautioned against the use of information from hospital record abstracts "for detailed research and evaluation."

This article draws its examples from our experience with the Manitoba Health Services Commission data, a comprehensive data bank generally illustrative of systems designed to keep track of patient contacts with providers. This study concentrates on information concerning diagnostic and surgical procedures, identifying those areas in which problems exist and those in which the data can be relied upon.

Following the orientation of Sechrest and Phillips,[5] researchers need to learn the extent to which various sources of bias operate in different kinds of records. As other investigators have mentioned,[6,7] positive results tend to be reported more frequently than negative results. Such an unfortunate bias makes it difficult to build on the experience of others. Knowing what does not work may be as important as knowing what does. Because large data banks seem likely to receive increasing research attention in the future, this discussion of biases in one data bank has wider implications.

Data banks often permit making reliability checks relatively easily. Most re-

From the Departments of Business Administration and Social and Preventive Medicine, University of Manitoba, Winnipeg, Manitoba.

This research has been supported by National Health Research and Development Projects Nos. 6607-453-44, 6607-1197-44 and 6607-1156-46, as well as by Research Scholar Award No. 606-1114-48 (to L.L. Roos) and Career Scientist Award No. 607-1001-22 (to N.P. Roos).

Requests for reprints: Leslie L. Roos, Jr., Ph.D., Faculty of Administrative Studies, University of Manitoba, Winnipeg, Manitoba R3T 2N2, Canada.

0025-7079/82/0300/0266/$01.05 © J. B. Lippincott Co.

search projects will not be able to afford the large-scale, independent checks performed in the reliability assessment of data collected in the National Hospital Discharge Survey.[1,8] Researchers using data banks need to learn how to perform relatively simple cost-effective checks on data quality. Some of the techniques in this article depend on the availability of a comprehensive data bank, permitting the merging of information collected independently from hospital and physician claims. Such data banks are available in several Canadian provinces.[9,10] They are relatively scarce in the United States, although some investigators have successfully utilized several types of hospital files (Medicare and PAS) for reliability studies.[1] Other techniques suggested here are applicable when just one type of file, either hospital or medical claims, is available.

Comparing Canadian with American information provides a sense of the generalizability of the Manitoba findings to health care data banks elsewhere. Such work both reinforces the applicability of the low-cost techniques described here and suggests ways in which these methods can be expanded.

Data Base

In Manitoba, several separate files—the registration, hospital and medical files—are maintained for payment and control purposes by the Health Services Commission (MHSC). This health insurance data base is characterized by universal coverage (all individuals registered in Manitoba, no matter where care is received). Hospital and medical care are documented in considerable detail because of the fee-for-service payment scheme.

The registration file contains data on the population covered, organized by family registration numbers. With minor exceptions, individuals can be uniquely specified by using birth date, sex and initial identifiers in addition to family registration number.* The medical claims file is the vehicle for paying physicians. Claims are submitted by physicians for services rendered to patients. These claims are filed on a temporal basis and contain the physician's identification number, the patient's identifying information, the patient's residence (checked against the population registry), tariff (determined by whether the visit is a consultation, complete physician examination, regional examination, follow-up visit, and the like), and diagnosis (ICD-8 code). If the visit is a consultation, the number of the referring physician is recorded. As described elsewhere, the claims appear to provide valid data on total patient–physician contact.[11] In a fee-for-service system, claims should not underestimate the volume of physician contact, and, since physician practice profiles are routinely audited with detailed analyses of outlier practice patterns, it is unlikely that a significant amount of overbilling exists.[12]

The hospital file is built on the basis of each admission and contains patient identification, dates of admission and discharge, limited information on services rendered, and one or more diagnoses. Surgical procedures are noted, as well as identification numbers of the surgeon or attending physician, consultants and anesthetist.

The data base generally meets the criteria for a health statistics information system forwarded by White.[13] Although each of these files is maintained separately, with no routine record integration, patient identifiers and physician numbers are unique across all files. Files can be built on individuals (all instances of care received from various physicians and in various hospitals over time), physicians (all care rendered over time), or hospitals (all admissions over time).

* The only systemic limitation on identifying individuals concerns two children in a family born in the same year, of the same sex, with the same initials. Claims for such individuals cannot be distinguished until one becomes 19 or marries.

Access to specified portions of the data base has been provided after approval by both the College of Physicians and Surgeons and the Manitoba Health Services Commission. Requests for data preparation are put together by the researchers and Commission staff, meeting to work out problems before the request is formalized. Confidentiality has been maintained through the use of abbreviated identifiers; patient name or address is never specified on the researchers' data. Prior to submission for publication, research papers are reviewed by Commission staff to ensure factual accuracy and to verify that individual patients or physicians cannot in any way be identified from the write-up.

Face Sheet Information

Maximum reliability might be expected when the information being sought is unambiguous and not subject to professional disagreement. As the Institute of Medicine study[8] suggested, "face sheet" data collected as part of a government registry seem likely to be of generally high quality in industralized countries. Information on address and marital status may become outdated, but reported date of birth should remain constant. In Manitoba, date of birth information is obtained for the master registry file and reliably duplicated on the claims file. In a check using survey data from 1971 interviews with over 4,500 elderly,[14] age reported by the interviewee and age calculated from the claims information proved identical in 92 per cent of the cases. Since the former was validated by official documents (birth certificate, citizenship papers, etc.) whenever possible, the reliability is encouragingly high.

Data on major events that are routinely collected in health care institutions also tend to be correctly captured in claims systems.[15] One hundred patients recorded as having died in hospital were identified from hospital claims; searches of vital statistics records corroborated these deaths. Ninety-eight per cent of the deaths (50) in Winnipeg hospitals and all of the deaths (50) occurring in rural hospitals were found in both sources. Such accuracy is dependent on the appropriate reporting system. Thus, pilot studies starting with vital statistics records showed no problems when comparisons were made with deaths reported on hospital claims, but some deaths (both in-hospital and out-of-hospital) were not recorded in timely fashion on MHSC's population registry before 1978. Only when appropriate organizational ties were set up with the Office of Vital Statistics did the reporting of such deaths on the registry become routinized.

Procedures and Diagnoses

Other research problems involve subject matter more ambiguous than births and deaths. Studies of elective surgery must consider the reliability associated with both procedures and diagnoses. The presence or absence of certain operative procedures (tonsillectomy, cholecystectomy, etc.) typically defines operations and control groups in longitudinal outcome studies. Inappropriate selection of diagnosis by the physician may affect the outcome analysis comparing these groups. The situation is complicated, since misclassification may increase or reduce the magnitude of association.[16]

Using several independent sources of information is basic for reliability studies. One type of check looks at procedures and diagnoses across two files, comparing those on the hospital claim with their counterparts on the operating physician's medical claim. Comparing the data on claims submitted by operating physician, anesthetist and assisting physician provides another way to invesigate reliability. In this later case, only one file (the medical claims file) is involved.

Reliability of Procedures

Surgical procedures are reasonably unambiguous events reported on the claims form submitted by the hospital to the Manitoba Health Services Commission for each patient stay. Tables 1 and 2 show general correspondence between hospital claims and surgeons' medical claims. With 722 cases overall, Table 1 contains relatively small numbers in cells suggesting an error. In only 12 cases did the MHSC tariff indicate an abdominal hysterectomy (total) while the ICD-8 operation code noted vaginal hysterectomy (total). The situation was reversed in just eight cases: MHSC tariff indicating vaginal hysterectomy with ICD-8 code noting abdominal. A similar table for gallbladder operations showed 246 cases in which both tariff and ICD-8 code agreed on a cholecystectomy; in 13 cases there was disagreement as to whether or not the common bile duct was explored.

Checks comparing tariffs filed by surgeons, anesthetists and assisting physicians had similarly high levels of agreement. As shown by Table 2, identical tariffs were filed in 91 per cent of the cases involving hysterectomy and in 89 per cent of the cholecystectomy cases. Only a single claim was filed in four and five per cent of the cases, respectively. Thus, direct discrepancies were found in only five (hysterectomy) and seven (cholecystectomy) per cent of the cases. In summary, the MHSC data appear highly reliable in recording the common surgical procedures of cholecystectomy and hysterectomy.

Comparisons of the Manitoba reliabilities with those found in the Institute of Medicine analysis of the National Health Discharge Survey abstracts should take note of the purpose underlying the data banks. As emphasized by a reviewer of this paper, the National Health Discharge Survey seeks to identify a single, principle procedure and/or diagnosis "for tabulation

TABLE 1. 1974 Hysterectomies
ICD-8 Operation Codes on Hospital Claims
and MHSC Tariffs on Surgeon's Medical Claims

ICD-8 Operation Code	4617 Abdominal Hysterectomy (total)	4631 Vaginal Hysterectomy	4621 Hysterectomy (subtotal)	4627 Hysterectomy (radical)	Total	Per Cent
69.1 abdominal (partial)	1	—	5	—	6	1
69.2 abdominal (total)	244	8	2	1	255	35
69.3 abdominal (radical)	—	—	—	4	4	1
69.4 vaginal (total)	12	141	—	—	153	21
67.0–67.9 operations + on 69.2 ovary	288	1	1	2	292	40
68.0–68.9 operations + on 69.2 fallopian tubes	11	1	—	—	12	2
TOTAL	556	151	8	7	722*	100
Per cent	77	21	1	1	100	

* Surgeon's medical claim was missing in 20 cases.

and later derivation of length of stay and discharge rates." The Manitoba data can be used for this purpose, but contain additional information (up to three diagnoses and three operative procedures) to facilitate other administrative and research tasks. The reliability of the Manitoba data may be somewhat higher than that from the Institute of Medicine study, but the methodological differences do not permit a definitive statement. The Institute research reported 85 per cent agreement between National Health Discharge Survey abstract and the reabstract for abdominal surgery, with 76 per cent agreement for gynecological surgery when procedures were coded to two digits. The National Health Discharge Survey's coders had some difficulty in deciding which of several related procedures was to be noted as the principal procedure.†

Reliability of Diagnoses

Several investigators have emphasized that diagnostic reliability approaching 100 per cent cannot be achieved. Koran's research noted substantial disagreement among physicians,[17,18] whether they were eliciting physical signs, interpreting diagnostic tests, making a diagnosis, recommending treatment or evaluating quality of care. Although the diagnosis of some conditions is undoubtedly more reliable than that of others,[19] 60 and 70 per cent levels of agreement are not unusual, even when physicians know that their clinical judgment is being examined.

The reliability of diagnoses appearing on hospital and medical claims can be assessed in several ways. First of all, independently collected information from hospital records can be compared with the MHSC data. Table 3 presents diagnoses among individuals aged 59 years and over who were noted on the hospital records as dying of myocardial infarction in 1974–1976. Diagnoses on the MHSC claims show a high level of agreement with those on the hospital record for four Winnipeg and four rural hospitals. Diagnoses indicating serious coronary problems occur in almost all of the individuals (152 out of 154). In similar fashion, a set of checks using the Manitoba Health Services Commission data and records from one Winnipeg hospital found 37 (95 per cent) of the diagnoses for gallbladder patients to be identical. Levels of agreement using the three-digit ICD codes were higher than those re-

† Ignoring ordering discrepancies on hospital discharge claims greatly increases levels of reliability.[4] Removing the assumption that the ordering of the diagnoses contributes to error raises the lowest reliability of any of the 15 Institute of Medicine diagnoses reviewed from 36.8% to 70.6%, with the overall reliability of diagnoses moving from 71.7% to 87.2%. Such high reliability implies that the National Hospital Discharge Survey should prove to be a fruitful research source.

TABLE 2. Comparison of Tariffs on Medical Claims Filed by Surgeon, Anesthetist and Assisting Physician

	Hysterectomy* (N = 742) Per Cent	Gallbladder and Biliary Tract Surgery† (N = 289) Per Cent
All tariffs identical	91	89
Only 1 claim filed	4	5
1 of 2 or more claims filed has discrepant tariff	5	7

* All hysterectomies in MHSC medical claims files for the period April 1–August 31, 1974.
† A 25 per cent sample of gallbladder and biliary tract surgery in MHSC medical claims files for the period April 1–August 31, 1974.

corded for the same conditions in an American study using independently recorded MEDPAR and PAS diagnoses.[3] Such findings suggest that, with appropriate caution, serious operations and life-threatening conditions can be meaningfully investigated using the MHSC data bank.

Time-Sequenced Relationships

Another technique for assessing claims data involves identifying logically time-sequenced relationships. Certain diagnoses should be associated with certain procedures. Thus, tonsillitis diagnoses were found in the year before tonsillectomy, while those of otitis media preceded adenoidectomy.[20] Several such checks were performed in analyzing the risks and benefits of hysterectomy.[21] For example, neither physician visits for pregnancy nor those for cervical disease should logically occur following a total hysterectomy. In fact, we found a 1–2 per cent error rate in the recording of these diagnoses. Other patterns, however, confirmed the general validity of the claims data: for example, menopausal symptoms increased after hysterectomy but not after cholecystectomy.

Intraphysician and Interphysician Agreement

Although recording and logical errors do not appear to present major problems with the hospital data, assessing reliability involves more than checking for simple errors. Psychometricians[22] have traditionally estimated the reliability of measures through the use of the same, or parallel, forms at two closely related time points. As noted in our previous work,[11] this suggests specific questions relevant to the Health Services Commission data:

1. Would the same informed observer (such as a physician) make the same diagnostic judgment twice when observing the phenomenon?

2. Would two different informed observers (such as two physicians) each make the same diagnostic judgment when observing the phenomenon?

These questions were answered using 1974 claims for gallbladder, biliary tract and related problems from patients having cholecystectomies in 1974. All such claims from physician visits taking place within the same 14-day period were used. Sensitivity checks showed only slightly improved results when a shorter (7-day) period between claims was chosen. Table 4 was produced by comparing the first and second claims within the 14-day period. The levels were defined in terms of their implications for gallbladder/biliary tract surgery. Level 1 diagnoses would be taken by many physicians as indicators for surgery. These included ICD-8 574 (chole-

TABLE 3. Comparing MHSC Diagnostic Data with Hospital Records

	Winnipeg (N = 75) Per Cent	Rural (N = 79) Per Cent
Identical MI (myocardial infarction) diagnoses in both MHSC and hospital record (ICD-8 diagnosis of 410)	92	89
MI diagnosis in hospital and related diagnosis* in MHSC record	5	5
MI diagnosis in MHSC and related diagnosis in hospital record	3	4
MI in hospital, unrelated diagnoses in MHSC record	—	3
MI in MHSC, unrelated diagnoses in hospital record	—	—

* Related ICD-8 Diagnoses:
411: Other acute and subacute forms of ischemic heart disease.
412: Chronic ischemic heart disease.
427: Symptomatic heart disease.
413: Angina pectoris.
414: Asymptomatic ischemic heart disease.
428: Other myocardial insufficiency.
429: Ill-defined heart disease.

The sample was made up of individuals aged 59 and over who were discharged with myocardial infarction diagnoses in Winnipeg and rural hospitals in 1974–1976.

TABLE 4. Intraphysician and Interphysician Reliability of Claims for Gallbladder, Biliary Tract and Related Diagnoses

	Intraphysician—Same Physician, 2 Observations Per Cent	Interphysician—Different Physicians, 2 Observations Per Cent
Diagnoses identical or on same level	80	58
Diagnoses one level apart	16	39
Diagnoses two levels apart	4	3
N of pairs	383	319
N of individuals	254	247
N of physicians	182	215, 168

The five cases of pancreatitis (ICD-8 577) were excluded from the analysis.

lithiasis), 575 (cholecystitis, cholangitis), 576 (other disease of gallbladder and biliary tract), 156 (carcinoma of gallbladder and bile ducts—just one case in this sample), and 157 (carcinoma of pancreas—no cases in the sample). Level 2 was made up of the common 785 diagnosis (symptoms referable to abdomen and lower gastrointestinal tract); this diagnosis is widely found both before and after cholecystectomy.[23] Level 3 was a catchall category grouping a number of codes dealing with diseases of the digestive system.

The most reliable single code with a substantial N was ICD-8 574 (cholelithiasis), with 84 per cent intraphysician and 59 per cent interphysician agreement. ICD-8 code 785 and the codes grouped into the "miscellaneous disease of the digestive system" category are responsible for much of the drop in reliability between intraphysician and interphysician comparisons. These two categories were unreliable in themselves, with considerable shifting back and forth between categories. There was some tendency for a less specific diagnosis to be given by the first physician seen, and a more specific judgment (ICD-8 574, 575 or 576) made by the second physician. This is as might be expected in a referral situation. Eliminating claims when one diagnosis was in the miscellaneous disease category raised percentage agreement by 14 per cent for the interphysician category.

These data, in conjunction with respiratory data discussed elsewhere,[11] emphasize the need for grouping three-digit ICD-8 diagnostic information from medical claims. Overreliance on particular individual diagnoses is clearly dangerous, given the instability of the processes underlying clinical judgment. Performing such checks early in the research process is recommended, since the reliability of diagnoses varies according to the organ system and the particular medical problem.[8] These types of checks provide valuable information both on how to group diagnoses and on the extent to which such diagnoses should be used in any specific research project.‡

Problems with the Data

Focusing on Certain Procedures

The reliability of some types of diagnoses and operative procedures may be particularly suspect. For example, determining whether or not surgery deprived a

‡ Under some circumstances, knowing which diagnoses go into which categories may be difficult. Any particular decision may take advantage of chance fluctuations in the data base being examined. In such situations, techniques of cross validation developed in other connections may be useful.[22]

woman of ovarian function was relevant for the hysterectomy research. Such a distinction was not possible using surgical tariffs, but ICD codes distinguished between the removal of one ovary and the removal of both ovaries (or of the one remaining ovary). Since ovaries were sometimes removed independently of hysterectomy, data from all women (N = 39) having an ovary taken out in the two years before or after the hysterectomy were checked to see if the coding of ovary removal was logically consistent. In the two cases in which one ovary was taken out following a prior removal, this second operation was coded as the removal of one ovary (ICD-8 67.2) rather than the removal of the remaining ovary (67.4). The diagnostic data did not permit making this critical distinction.

Another surgical study deals with the risks and benefits associated with hip arthroplasty (87.1). A comparison of the procedures coded 87.1 with information written on the admission/separation form filed by the hospital suggested that almost 50 per cent of the procedures coded 87.1 should be excluded from the study (Thompson prosthesis, Mueller plates and pins, etc.). Distinguishing among procedures may be possible using the physician's tariff filed, since payment tariffs are often more specific than operation codes. The problem here is not strictly one of reliability, but rather of not accepting potentially ambiguous categories at face value.

In an effort to assess reliability along the lines of the single-file analysis of surgical procedures, diagnoses on medical claims from surgeons, anesthetists and assisting physicians were compared. Table 6 uses the diagnostic categories outlined in Table 5, highlighting the lack of agreement among diagnoses recorded by physicians involved in hysterectomies. Similar results were found for gallbladder surgery. Twenty-four per cent of the cases showed complete agreement among the different physicians, while 68 per cent concerned level 1 diagnoses pertaining to gallbladder or biliary tract diseases. Anesthetists seldom discriminated among diagnoses within the broad gynecologic and gallbladder-related categories. Most anesthetists merely selected a single diagnosis (such as 574, cholelithiasis), repeating this whenever it might be relevant (i.e., for a cholecystectomy).

Such information suggests that the reliability of a diagnosis might depend on such factors as: 1) the amount of time spent on selecting the diagnosis and entering this information; and 2) the closeness of the physician (both psychologically and technically) to the patient's problem. Given the nature of an anesthetist's work, the physician giving the anesthetic often need not be concerned about the specific diagnosis.

Underreporting

Funding permitted small-scale use of the "reabstracted record" method to check on the reliability of other information collected by the Manitoba Health Services Commission.[24] The reporting of in-hospital consultations and some procedures from the hospital to MHSC is less consistent than the reporting of major operations and

TABLE 5. ICD-8 Diagnostic Categories Used in Gynecologic Research

Gynecologic Diagnoses	
630–677, Y06	Pregnancy
180–184	Malignant neoplasm
091, 098, 612–614, 616, 622	Sepsis
218–221, 615	Benign neoplasm
623, 624, 878	Acquired abnormal anatomy
625	Endometriosis
Y09	Prophylactic sterilization
620, 621, 234	Cervical disease
626	Menstrual disorders
099, 256, 595, 627–629, 752, 786, 939	Other gynecologic disorders

TABLE 6. Comparison of Diagnoses on Medical Claims Filed by Surgeon, Anesthetist and Assistant Physician for Hysterectomy

	Per Cent
All are gynecologic diagnoses in same category*	13
All are gynecologic diagnoses but not in same category	79
Gynecologic and nongynecologic diagnoses	3
All are nongynecologic diagnoses	0
Only one claim filed	4
N = 742	

* Gynecologic diagnosis categories are as described in Table 5; any three-digit code not in that list constitutes a nongynecologic diagnosis.

of diagnoses. The form submitted by the hospital to the Health Services Commission is acceptable when consultations and certain procedures are not indicated. On the other hand, diagnoses must be specified on the form; leaving out surgery would be a major error. Table 7 shows an underreporting of consultations on MHSC claims from both urban and rural hospitals.

Table 8 illustrates a similar underreporting of various procedures on the MHSC records of these cardiac patients. This underreporting is important because relatively few patients undergo such procedures in the first place. Procedures noted on the MHSC record were almost always performed, but the absence of an entry does not mean the procedure was not done. In addition to the procedures shown in Table 8, radiotherapy, shock therapy, hemodialysis, myelography, single blood transfusion, chemotherapy and hyperalimentation were all checked, but these procedures were not frequent enough to document difficulties in the recording. Finally, a check using 39 gallbladder patients from one Winnipeg hospital found a similar underestimation of the amount of cholecystography performed.

Why do reliability problems occur in some cases but not in others? Both technical factors, such as whether an item had to be noted on the MHSC claim, and substantive factors may be involved. Surgical procedures are both recorded in a central position on the hospital claims form and clearly of major medical importance. On the other hand, the miscellaneous procedures are to be checked off toward the bottom of the form and are not required items.

TABLE 7. Comparing MHSC Consultation Data with Hospital Records

	MHSC and Hospital Record			
	Both Yes	Both No	Yes—MHSC No—Hospital Record	No—MHSC Yes—Hospital Record
In-Hospital Consultations—Winnipeg (N = 103 records)				
First consultation	55%	30	—	15
Second consultation	14%	71	3	13
In-Hospital Consultations—Rural (N = 124 records)				
First consultation	12%	82	1	5
Second consultation	2%	98	—	1

Percentages are calculated horizontally.
The sample was made up of individuals aged 59 and over who were discharged with myocardial infarction diagnoses in Winnipeg and rural hospitals in 1974–1976. This table includes data from multiple hospitalizations of the same patient, resulting in a larger N than that of Table 3.

TABLE 8. Comparing MHSC Miscellaneous Procedures Data with Hospital Records

	MHSC and Hospital Record			
	Both Yes	Both No	Yes—MHSC No—Hospital Record	No—MHSC Yes—Hospital Record
Physiotherapy	9%	83	3	4
Inhalation therapy	7%	85	—	8
EKG	98%	—	—	2
Multiple blood transfusion	3%	90	1	6
Cardiac monitoring	69%	10	—	21

Percentages are calculated horizontally; 103 individual records from Winnipeg hospitals were used. The sample was made up of individuals aged 59 and over who were discharged with myocardial infarction diagnoses in Winnipeg and rural hospitals in 1974–1976. This table includes data from multiple hospitalizations of the same patient, resulting in a larger N than that of Table 3.

Discussion

Several general themes emerge from this review of the quality of the Manitoba data. Finding information recorded independently by separate individuals or organizations, at two different times, or in two or more data files is one key to performing reliability studies inexpensively. Secondly, when a particular event has certain logical implications vis-à-vis another, inconsistencies can be located and their importance assessed.

We have adopted a practical outlook to problems of reliability, emphasizing the sorts of checks that can be done with minimal resources. Although some checks using primary data collection (i.e., going back to the physical hospital records) have been mentioned, computerized approaches have been emphasized. In the process of organizing claims for data analysis, many of these checks can be performed using such packages as SPSS with comparatively little additional effort. The general similarity between the Manitoba data and those from the Institute of Medicine study points up the usefulness of the computer-based techniques described here. Any redundancy—similar information in one or several data files—may permit reliability checks. Thus, Wennberg and Gittlesohn's[3] checks using PAS and Medicare files produced results convergent with the IOM and Manitoba data discussed above—face sheet data and discharge status were seen as the most reliable, procedure data as next in quality, and diagnostic information as presenting the most problems.

Although acceptable reliability of the information contained in large data bases is a prerequisite for meaningful research, researchers need to learn other things about such data banks. Since experimental and quasi-experimental designs require longitudinal information, problems of following individuals through time must be addressed. Tracing individual calls for attention to the reliability and completeness of identifiers. Mobility and life cycle changes complicate the process. Future papers will address this important problem of building individual histories.[14]

Acknowledgments

The authors would like to thank the Manitoba Health Services Commission for facilitating access to claims data. Interpretations and viewpoints contained in this article are the authors' own and do not necessarily represent the official opinion of either the Manitoba Health Services Commission or Health and Welfare, Canada.

References

1. Brook RH, Williams KN. Evaluation of the New Mexico peer review system 1971–1973. Med Care 14(Suppl)1976;1.

2. Federspiel CF, Ray WA, Schaffner W. Medicaid records as a valid data source: the Tennessee experience. Med Care 1976;14:166.

3. Wennberg JE, Gittelsohn AM. A small area approach to the analysis of health system performance. Hyattsville, Md.: Health Resources Administration, 1980. (Final report, HRA contract no. 291-76-0003.)

4. Demlo LK, Campbell PM, Brown SS. Reliability of information abstracted from patients' medical records. Med Care 1978;16:995.

5. Sechrest L, Phillips M. Unobtrusive measures: an overview. In: Sechrest L, ed. New Directions for Methodology of Behavioral Science: Unobtrusive Measurement Today. San Francisco: Jossey-Bass, 1979.

6. Peto R, Pike MC, Armitage P, et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design. Br J Cancer 1976;34:585.

7. Rosenthal R. The "file drawer problem" and tolerance for null results. Psychol Bull 1979;86:638.

8. Institute of Medicine. Reliability of National Hospital Discharge Survey Data. Washington: National Academy of Sciences, 1980.

9. Beck RG, Horne JM. Economic class and risk avoidance: experience under public medical care insurance. J Risk and Ins 1976;43:73.

10. Dyke FJ, Murphy FA, Murphy JK, et al. Effect of surveillance on the number of hysterectomies in the province of Saskatchewan. N Engl J Med 1977;296:1326.

11. Roos LL, Nicol JP, Johnson CF, et al. Using administrative data banks for research and evaluation. Eval Q 1979;3:23.

12. Roos NP. Impact of the organization of practice on quality of care and physician productivity. Med Care 1980;18:347.

13. White KL. Information for health care: an epidemiological perspective. Inquiry 1980;17:296.

14. Roos LL, Nicol JP. Building individual histories with registries: a case study. Winnipeg: Faculty of Administrative Studies, University of Manitoba, 1981.

15. Gittelsohn AM, Senning J. Studies on the reliability of vital and health records: comparison of cause of death and hospital record diagnoses. Am J Public Health 1979;69:680.

16. Fleiss JL. Statistical Methods for Rates and Proportions. New York: John Wiley & Sons, 1973.

17. Koran LM. The reliability of clinical methods, data and judgments. N Engl J Med 1975;239:642.

18. Rutkow IM, Gittelsohn AM, Zuidema GD. Surgical decision making: the reliability of clinical judgment. Ann Surg 1979;190:409.

19. Engel L, Davis BJ. Medical diagnosis: present, past, and future. Arch Intern Med 1963;112:512.

20. Roos NP, Henteleff PD, Roos LL. A new audit procedure applied to an old question: is the frequency of t and a justified? Med Care 1977;15:1.

21. Roos NP. Hysterectomies in one Canadian province: a new look at risks and benefits. Winnipeg: Department of Social and Preventive Medicine, Faculty of Medicine, University of Manitoba, 1981.

22. Campbell JP. Psychometric theory. In: Dunnette MD, ed. Handbook of Industrial and Organizational Psychology. Chicago: Rand, McNally, 1976;185–222.

23. Horne JM, Beck RG. Temporal patterns in the use of health services leading to cholecystectomy: a process evaluation using insurance records. Med Care 1978;16:1006.

24. Goldberg J, Gelfand HM, Levy PS. Registry evaluation methods: a review and case study. In: Sartwell PE, Nathanson N, eds. Epidemiologic Reviews, vol. 2. Baltimore: Johns Hopkins, 1980;210–220.

MEDICAL CARE
March 1982, Vol. XX, No. 3

The Interaction Among the Criteria Physicians Use When Prescribing

ROBERT N. ZELNIO, PH.D.

A randomly selected sample of Iowa physicians was surveyed using a self-administered mail questionnaire that measured the relative importance of prescribing criteria as well as selected physician and practice characteristics. Based on a response rate of 68.3 per cent, it was found that the physicians' criteria for selection of a drug to prescribe could be reduced to two groups. The first group is used to select the generic drug or therapeutic class and is of primary concern to the physicians. Those criteria used in specifying a brand of drug product are of lesser importance to physicians and constitute the second group. It was further discovered that physicians could be segregated into two distinct segments based on the importance of the prescribing criteria. It was not possible, however, to explain the composition of the groups using the physician and practice characteristics measured.

THE DEMAND FOR PHARMACEUTICALS has been described as derived demand.[1] This description aptly conveys the fact that the market for drug products is unique in that the ultimate user or consumer of the product is not the primary decision maker in the purchase process. Rather, this process is controlled by the patient's physician, who acts as an agent in deciding whether or not a purchase will be made and, if so, which product will be purchased. The reality of derived demand in combination with the fear that drug therapy is often inappropriate or employed to excess has prompted many studies of physicians' prescribing habits.

While the suggestion has been made that appropriateness, effectiveness, safety and economy are the four criteria that define rational prescribing,[2] research has revealed that the prescribing process is more complex and often less rational. Prescribing would appear to be a function of five broad factors: 1) the patient's disorder; 2) the product; 3) the manufacturer; 4) the physician; and 5) the physician's environment. Assumedly because its impact is obvious and important, the patient's disorder has not been the subject of studies of the prescribing process per se.

The product characteristics studied center on the drug's effectiveness, risk, cost and pharmaceutical properties. In one form or another, a drug's effectiveness in a given situation has been found to be the most important reason for its being prescribed.[3-5] The risk associated with a drug product is usually studied by means of its potential side effects and has been said to be second to effectiveness in influencing prescribing.[3,4] In a review of relevant research, it was concluded that physicians do not consider costs when prescribing.[6] Nevertheless, the physicians sampled in such studies often report that cost is an important consideration.[4,7] The pharmaceutical properties cited as influencing

From the College of Pharmacy, University of Iowa, Iowa City, Iowa.

This research was supported by an NIH Biomedical Research Grant.

Reprint requests: Robert N. Zelnio, Ph.D., Assistant Professor, University of Iowa, College of Pharmacy, Iowa City, IA 52242.

0025-7079/82/0300/0277/$00.95 © J. B. Lippincott Co.

physicians' choice of drugs include the ease with which the drug can be administered and the availability of various dosage forms.[3,5]

The impact of drug manufacturers has been studied from two perspectives. First, it has been shown that the company's image or reputation is associated with physicians' preference among drug products.[8,9] Second, manufacturers have been studied as sources of information about drug products. The information provided by manufacturers through detail men, journal advertising, direct mail advertising, sampling, the Physicians' Desk Reference, at meetings or conventions, and through continuing education materials has been said to be highly influential.[3,7,9–13] While all such sources of drug information appear to be effective in influencing physicians, it should be noted that most studies point to detailing as most effective, followed by journal and direct mail advertising. Furthermore, the impact of these sources seems to diminish as physicians gain experience with drug products.

Education is a physician characteristic that has been the subject of repeated study and has been related to the prescribing process in the form of medical school attended, educational qualifications, year of graduation and extent of postgraduate education.[4,5,10,14,15] The consensus appears to be that more recent graduates with more postgraduate education are better prescribers. Prescribing has also been related to physician age, sex, experience, attitudes, specialty and level of job satisfaction.[4,5,9,10,16,17] The evidence regarding the effect of these characteristics on prescribing is, however, often conflicting, which is partly the result of the fact that some studies relate the characteristics to a measure of prescribing quality while others use the number of prescriptions written.

Noncommercial sources of information and the physician's practice characteristics are two categories of environmental variables that have been examined for their effect on prescribing. In general, noncommercial sources of information, of of which journal articles and colleagues are the most important, tend to be more influential as physicians gain experience with drug products and begin to adopt them on a regular basis.[3,11,18–20] Journal articles are often said to be the second most important source of information and serve to legitimize or reinforce physicians' decisions.[3,4,9,11,12] In addition to increasing over time, the impact of articles appears to be enhanced as physicians read more journals and become more accepted in their communities.[4,12]

The examples and testimonials of peers seem to be the method by which colleagues influence prescribing.[21] Their impact is supported by the finding that the rate of new product adoption is enhanced by the existence of social and professional links between physicians.[18] But while the impact of physicians on their colleagues has been established, other health care colleagues apparently exert little influence on prescribing. For instance, the few studies that have addressed the impact of the pharmacist report that although physicians frequently consult with pharmacists, their influence is weak and may depend on the type of information sought and the size of the city in which physicians practice.[3,7,22] It should be noted, however, that the study of the effect of ancillary personnel on prescribing is sparse and inconclusive, and predates much of the expansion of the traditional pharmacist and nurse roles, not to mention the creation of new paraprofessionals. Yet despite this fact, one study has associated the increased use of ancillary personnel with a measure of better prescribing.[9]

The organization of the physician's practice is one variable that combines the effect of colleagues with other practice characteristics. The opportunity to develop or take advantage of associations with other physicians affects the prescribing process.

Thus it has been found that group practices, hospital affiliations and private versus academic practices influence how a physician prescribes.[3,7,9,23] The physician's patient population is a second practice characteristic that has been found to influence prescribing. While the patients themselves are said to have a weak, although underestimated, impact on prescribing, it has been found that the size of the patient population is inversely associated with the number of prescriptions written and the quality of prescribing, and that physicians having patient populations with high incomes tend to prescribe newer and more expensive drug products.[4,5,9,12,15]

Although most if not all of the studies reviewed have examined aspects of the prescribing process in isolation, many have either concluded or implied that the variables relevant to this process do not always exert their effects independently. For instance, Smith suggests that journal advertisements probably affect prescribing in the presence of other variables.[13] Furthermore, it has been stated that the effects of various sources of information are dependent on the physician's experience with drug products, practice organization, community acceptance and degree of innovativeness.[3,4,7,11,12,18–20,24]

Despite the observation of these and other such interactions, the studies of the prescribing process have not simultaneously considered the relative importance of the various prescribing criteria, nor have their interdependencies been adequately studied. As a result, the objective of this study was to measure the importance of physicians' prescribing criteria relative to one another and to explore their interdependencies using multivariate analyses.

Methods

This objective was addressed by means of a mail survey of 250 Iowa physicians selected from the *AMA Directory*. The selection process was random with the stipulation that the sample contain only physicians who could be expected to prescribe drugs in the course of their practices. This requirement excluded practitioners with specialties such as radiology, anesthesiology and pathology. Each physician in the sample was sent a questionnaire, cover letter and a self-addressed, stamped return envelope. The survey instrument was designed and administered according to recommendations by Dillman.[25] Nonrespondents received a postcard reminder one week after the initial mailing urging them to complete and return the questionnaire. Replacement questionnaires and return envelopes were sent to the remaining nonrespondents after intervals of three and six weeks.

The importance of eight prescribing criteria was scaled using paired comparisons. These criteria were chosen from among those identified by previous studies. While there may be others or other ways to phrase them, it is believed that the eight criteria used in this investigation accurately represent the most important factors identified to date. The paired comparisons method presents all possible pairs of the criteria in randomized order and requires the respondent to select the most important one from each pair. Thus, each prescribing criterion appeared in seven pairs, for a total of 28 pairs and would have a scale value ranging from 0 to 7. The paired comparison method has the advantages of being easy for the respondent when the items to be rated number less that ten and of providing interval level data.[26] In addition to the paired comparisons, the questionnaire asked each physician for background information on their types of practice, size of practice, estimated patient income and age.

Results and Discussion

Of the 250 questionnaires sent to the physician sample, 23 were returned as un-

deliverable to the desired respondent for a variety of reasons. A total of 155 completed questionnaires were returned out of the remaining 227, for a total response rate of 68.3 per cent. However, only 118 respondents completed the entire set of paired comparisons, for a usable response rate of 52 per cent. The characteristics of the sample are displayed in Table 1. As can be seen, the sample is primarily composed of middle-aged physicians in private practice with large, middle-income patient populations. With the exception of type of practice, there are no comparable data enabling an examination of the representativeness of the sample depicted by the information of Table 1. In 1978, it was reported that 2,355 physicians in Iowa were office based, while 680 were hospital based. The remainder were in teaching, research, administration, retired or unclassified.[27] Thus, 77.6 per cent of the practicing physicians were office based and 22.4 per cent were hospital based. This compares with 76.2 and 23.9 per cent, respectively, when the clinic physicians of Table 1 are combined with those practicing in hospitals to yield a more comparable institutional classification. While this result provides evidence that the sample is representative of physicians in Iowa on the only variable for which outside information is available, a comparison based on a single variable should not be taken as proof of general representativeness.

The criteria used when evaluating drug products during the prescribing process are listed in Table 2 in order of importance

TABLE 1. Sample Characteristics

Characteristics	Frequency	
	n	Per Cent
Type of practice		
Private, solo, office based	49	36.6
Private, group, office based	53	39.6
Hospital based	20	14.9
Clinic based	12	9.0
Total	134	100.0
Size of practice		
Less than 500 patients	13	10.0
500 to 999 patients	11	8.5
1000 to 1999 patients	34	26.2
2000 to 2999 patients	28	21.5
3000 or more patients	44	33.8
Total	130	100.0
Estimated average annual gross family income of patient population		
Less than $10,000 per year	4	3.2
$10,000 to $19,999 per year	64	50.8
$20,000 to $29,999 per year	53	42.1
$30,000 to $39,999 per year	3	2.4
$40,000 or more per year	2	1.6
Total	126	100.1
Physician's age		
Less than 35 years old	10	7.4
35 to 44 years old	36	26.5
45 to 54 years old	37	27.2
55 to 64 years old	36	26.5
65 or older	17	12.5
Total	136	100.1

TABLE 2. Relative Importance of Criteria Used in the Prescribing Process

Prescribing Criteria	Frequency of Selection	Mean Importance	$t_{(117)}$*
Potential side effects	627	5.3	42.7
Probable efficacy	613	5.2	34.4
Contraindications	611	5.2	32.5
Dosage form availability	268	2.3	16.3
Drug Cost	264	2.2	14.2
Source of drug information	255	2.2	14.0
Frequency of administration	229	1.9	14.0
Manufacturer's reputation	228	1.9	12.1
n		118	

* Significantly different from 0 beyond the 0.05 level.

to the sample physicians along with the frequency and mean number of times each was selected from the paired comparisons. The higher the frequency and mean, the more important each criterion was relative to all others. As stated, the scores on which the means were computed range from 0 to 7, as each criterion appeared in seven pairs. The most important criteria deal with the safety, efficacy and appropriateness of therapy. Of lesser importance are criteria reflecting the mechanics of therapy such as dosage forms, cost and administration. In addition, while previous studies have found the manufacturer and information sources to be highly influential, this study found them to be of little importance. One explanation might be that the manufacturer was separated from the source of information, and since the dominant criteria concerned the safety, efficacy and appropriateness of therapy, perhaps physicians are influenced less by the manufacturer or source of information and more by the content or nature of the information. But it should be noted that each of the means of Table 2 was found to be significantly different from 0 using Student's t-test. This suggests that each of the criteria is of some importance to physicians.

Because the data point to the existence of two distinct groups of prescribing criteria, an F test of the hypothesis that all eight means are equal was performed. As shown in Table 3, this hypothesis can be rejected, lending credibility to the observation that some criteria are more important than others. Also displayed in Table 3 are the results of contrasting a linear combination of the means of the first three criteria with a linear combination of the remaining four. The fact that the difference was significant not only supports the existence of two groups of prescribing criteria but also

TABLE 3. Tests of the Mean Importance of Prescribing Criteria

Source	SS	df	MS	F*
Between groups	2,167.89	7	309.7	119.11
Within groups	2,433.84	936	2.6	
Total	4,601.73	943		

$$t_{(943)} = \frac{1/3(5.2 + 5.2 + 5.3) - 1/5(1.9 + 2.2 + 2.3 + 2.2)}{\sqrt{2.6\,(0.533/118)}}$$

$$= 28.8^*$$

* Significant beyond the 0.05 level.

shows that the first is significantly more important than the second.

The structure of these data was further explored using factor analysis. It was expected that the factor analysis would combine interdependent prescribing criteria along the two dimensions indentified above. But, as can be seen in Table 4, three factors were extracted that have eigenvalues of 1 or more and accounted for all the variance of the data. Although these results do not confirm the two-dimension hypothesis, three limitations of this analysis must be noted. First, there is only one variable that is clearly associated with the third factor. Second, the third factor accounts for a small portion of the total variance. And third, two variables do not load highly on one factor and low on the others as is desired. "Dosage form availability" loads moderately high on two factors, while "probable efficacy" does not load highly on any factor. With larger samples, these limitations might not occur, and a two-factor solution might still be likely. But, even with the limitations, this solution was superior to all others and provides some support for the grouping of criteria observed earlier.

The first factor might be named "brand attributes" because all the variables loading highly deal with individual drug product characteristics, which may vary from manufacturer to manufacturer. In contrast, the variables constituting the second factor relate to the safety, efficacy and appropriateness of therapy in general, and might be used in deciding on the therapeutic class of drug to prescribe. When "probable efficacy" is ignored for purposes of interpretation as is often recommended for variables not loading highly on any factor, the second factor can be named "therapeutic safety." The last factor can easily be named "drug cost." Because first and third factors are composed of variables that are of secondary importance to physicians, while those of factor 2 are most important, it would seem that the physicians in this sample are primarily concerned with the selection of the generic drug or therapeutic class, and only slightly concerned with the

TABLE 4. Factor Analysis of Prescribing Criteria*

Prescribing Criteria	Factors			Commonalities
	1	2	3	
Frequency of administration	0.70	0.08	0.02	0.49
Manufacturer's reputation	−0.50	−0.00	−0.10	0.26
Source of drug information	−0.54	−0.02	0.11	0.30
Dosage form availability	0.45	0.20	0.38	0.39
Contraindications	−0.19	0.83	−0.09	0.74
Probable efficacy	0.15	0.29	0.00	0.30
Potential side effects	0.06	0.54	0.07	0.30
Drug cost	0.21	0.04	−0.60	0.40
Eigenvalues	4.18	2.50	1.32	
Cummulative variance accounted for	0.52	0.84	1.00	

* After orthogonal rotation.

selection of a particular brand of drug product. Such a decision process would be consistent with the lack of importance of the drug manufacturer.

When the importance ratings of the prescribing criteria were cross-tabulated with the physician and practice characteristics of Table 1, no significant relationships were found. While this can be taken to mean that the importance of the prescribing criteria does not vary with the four variables measured in this study, it would not necessarily mean that the sample is homogeneous with respect to the criteria. As a result, cluster analysis was performed in order to investigate further whether or not the importance of the prescribing criteria varied across physicians. In fact, a two-cluster solution emerged when judged according to the change in the maximum distance within clusters and the ratio of the number of distances within clusters that are less than or equal to the maximum to all distances less than or equal to the maximum.

Table 5 shows the results of cross-tabulating the mean importance of prescribing criteria by cluster. The criteria are listed in order of importance to physicians in cluster 1. As can be seen, 53 physicians make up cluster 1, while 65 are in cluster 2. Two observations can be made immediately. First, the order of criteria within both clusters is different from that of the entire sample, as shown in Table 2. This supports the cluster solution to the extent that it is distinct from the unclustered solution. The second observation is that the listing of the criteria differs from one cluster to the other, with the mean importance ratings differing significantly in five cases. Thus, there are two segments of physicians, and while the drug prescribed is a function of the same criteria for each segment, the importance of those criteria varies from segment to segment. The physicians of cluster 1 are significantly more concerned with the probable efficacy, cost and frequency of administration of therapy, and significantly less concerned with the source of drug information and manufacturer's reputation. Both segments are similar to the extent that the safety, efficacy and appropriateness of therapy continue to be dominant, but the second segment values less objective criteria such as the source of drug information and manufacturer's reputation over more objective criteria related to the mechanics of therapy.

The implications of such clusters are many. For example, pharmaceutical manufacturers should seek to provide the type of information valued by physicians. Both clusters require information on the safety, efficacy and contraindications of therapeutic agents. But the segments differ to the extent that the first would prefer information on the mechanics of therapy, while the second would seem to prefer information on the manufacturer itself and its sources of data. For society, at large, the existence of clusters that imply differing prescribing processes would simply raise questions about the effects of the differing processes. For example, is the quality of therapy higher in one cluster or the the other? If so, strategies should be devised to improve the process of prescribing in the lower quality cluster. One strategy might be for

TABLE 5. Mean Importance of Criteria by Cluster

Prescribing Criteria	Cluster 1	Cluster 2
Probable efficacy	5.7	4.8*
Potential side effects	5.3	5.3
Contraindications	5.2	5.2
Drug cost	3.5	1.2*
Frequency of administration	2.7	1.4*
Dosage form availability	2.5	2.1
Source of drug information	1.5	2.7*
Manufacturer's reputation	1.3	2.5*
n	53	65

* Means are significantly different across clusters at the 0.05 level.

the physicians in such a cluster to delegate responsibility for those criteria that are of less importance to them.

Summary and Conclusions

In an effort to explore the relative importance of physicians' prescribing criteria and the interaction among them, a randomly selected sample of 250 Iowa physicians was surveyed with a self-administered, mail questionnaire. While the questionnaire itself measured physician and practice characterisites and the importance of each criterion relative to all others presented, multivariate analytical tools were used to explore their interdependencies. Although limited by the use of a small, geographically distinct sample, it is believed that this research is unique and contributes to the area of study in at least three ways.

First, no other study has examined the influence of prescribing criteria in a way that allows the investigation of their interdependencies. Previous studies have yielded rankings of criteria based on their individual ratings. Paired comparisons as well as factor and cluster analyses explicitly analyze the importance of each criterion relative to all others.

Second, while each of the criteria was of some importance, the mean importance ratings of the criteria as well as frequency of selection suggest that they might be summarized by two distinct groups. The first group is composed of those criteria that are most important to physicians and might be used to select the generic product or therapeutic class of drug product. The second group includes criteria of lesser importance that would be of use in prescribing a specific brand of drug product. Interestingly, this finding agrees with a report by Caplow and Raymond that physicians claim to select drugs because of their operational characteristics rather than brand, price or popularity.[3]

Third, while it was not possible to differentiate between physicians using the background variables measured in this study, two distinct clusters of physicians do exist based on the relative importance of their prescribing criteria. Most probably, the composition of these segments could be explained by other variables not measured in this research, such as specialty, year of graduation or board certification.

Because of the sample limitations identified, future research should be performed to replicate these findings with larger, more broadly based samples. Such research should be conducted using multivariate methods and should incorporate additional background variables that might be capable of explaining the composition of physician segments. Additionally, the implications of these findings on the prescribing process and marketing of drugs should be examined. Such implications might include the reasons why physicians are relatively unconcerned with brand selection, the possibility for interaction between physicians and pharmacists in the selection of a brand of drug, and whether or not the activities of pharmaceutical manufacturers have or should evolve in light of these findings.

References

1. Smith MC. When friends or patients ask about . . . the cost of drugs. JAMA 1976;235:294.

2. Parish P. Sociology of prescribing. Br Med Bull 1974;30:215.

3. Caplow T, Raymond JJ. Factors influencing the selection of pharmaceutical products. J Mktg 1954;19:20.

4. Lilja J. How physicians choose their drugs. Soc Sci Med 1976;10:364.

5. Miller RR. Prescribing habits of physicians: a review of studies on prescribing of drugs parts IV–VI. Drug Intell Clin Pharm 1973;7:558.

6. Zelnio RN., Gagnon JP. The effects of price information on physician prescribing patterns—a literature review. Drug Intell Clin Pharm 1979;13:158.

7. Lowy DR, Lowy L, Warner RS. A survey of physician's awareness of drug costs. J Med Educ 1972;47:349.

8. Bauer R. Risk handling in drug adoption: the role of company preference. Public Opinion Quarterly 1961;25:546.

9. Becker MH, Stolley PD, Lasagna L, et al. Differential education concerning therapeutics and resultant physician prescribing patterns. J Med Educ 1972;47:121.

10. Stolley PD, Lasagna L. Prescribing patterns of physician. J Chronic Dis 1969;22:397.

11. Miller RR. Prescribing habits of physicians: a review of studies on prescribing of drugs parts VII–VIII. Drug Intell Clin Pharm 1974;8:82.

12. Worthen DB. Prescribing influences: an overview. Br J Med Educ 1973;7:112.

13. Smith MC. Drug product advertising and prescribing: a review of the evidence. Am J Hosp Pharm 1977;34:1221.

14. Garner JD, Smith MC, Mikeal RL. A study of relations of physician specialty, medical school and prescribing patterns. Paper presented to the Academy of Pharmaceutical Sciences, Economic and Administrative Sciences Section, American Pharmaceutical Association, San Francisco, 1971;9.

15. Hemminki E. Review of literature on factors affecting drug prescribing. Soc Sci Med 1975;9:111.

16. Knoben JE, Wertheimer AI. Physician prescribing patterns—therapeutic categories and age considerations. Drug Intell Clin Pharm 1976;10:398.

17. Melville A. Job satisfaction in general practice: implications for prescribing. Soc Sci Med 1980;14A:498.

18. Coleman JS, Katz E, Menzel H. Medical innovation—a diffusion study. Indianapolis: Bobbs-Merrill, 1966.

19. Ferber R, Wales HG. The effectiveness of pharmaceutical advertising: a case study. Journal of Marketing 1958;23:395.

20. Miller RR. Prescribing habits of physicians: a review of studies on prescribing of drugs parts I–III. Drug Intell Clin Pharm 1973;7:496.

21. Cluff LE. The prescribing habits of physicians. Hosp Pract 1967;2:100.

22. Reis L. 81% of high prescribing GP's contact Rx men for new drug data. Am Prof Pharmacist 1968;34:34.

23. Muller C. Outpatient drug prescribing related to clinic utilization in four New York City hospitals. Health Serv Rev 1968;3:142.

24. Winick C. The diffusion of an innovation among physicians in a larger city. Sociometry 1961;24:384.

25. Dillman DA. Mail and telephone surveys: the total design method. New York, Wiley, 1978;187.

26. Hughes GD. Attitude measurement for marketing strategies. Glenview, Ill. Scott Foresman, and Co, 1971;103-04.

27. Wunderdorman LE. Physician Disturbance and Medical Licensure in the U.S., 1978. Monroe, Wis.: Center for Health Services Research and Development, American Association, 1979;108.

MEDICAL CARE
March 1982, Vol. XX, No. 3

Does Cost Information Availability Reduce Physician Test Usage?

A Randomized Clinical Trial With Unexpected Findings

DAVID I. COHEN, M.D., M.Sc., PAUL JONES, PH.D.,
BENJAMIN LITTENBERG, B.A., AND DUNCAN NEUHAUSER, PH.D.

Four similar teams of physicians associated with similar inpatient units and randomly assigned patients were used to study the effect of providing physicians with cost information about their use of lab tests and x-rays. Two teams received information about lab test costs, and two teams received x-ray test costs. Test usage fell during the experimental conditions and continued to fall after the experimental period ended in teams in which there was an interested leader.

BECAUSE PHYSICIANS have the power to admit to hospitals, order tests and procedures, prescribe medications, and to influence repeat ambulatory care visits, physicians are believed to control 80 per cent of the costs of medical care. Of the $192 billion spent on medical care in 1978, $76 billion went to hospital care.[1] As much as 25 per cent of the average patient's hospital bill will be for laboratory tests.[2]

We hypothesized that providing physicians with information about the costs of their testing would lead to a reduction in test usage but that this reduction would diminish when the feedback of information ceased. Four internal medicine teams (firms) were provided with feedback over a four-week interval. The sharpest drop in laboratory test usage occurred in the two firms that did *not* receive laboratory test cost data. The results were not consistent with our initial hypotheses, although after the fact they are understandable.

The Firm Structure

Cleveland Metropolitan General Hospital is a 600-bed general and tertiary care hospital, with a major teaching affiliation with Case Western Reserve University, School of Medicine. The general internal medicine service is divided into four separate units called "firms," each consisting of a team of residents and attending physicians responsible for an identical 28-bed inpatient unit, outpatient clinics and similar patient populations based on their random assignment to the four firms. The medical residents stay with their firm throughout their three years, rotating off the firm for subspecialty experience during their second year. Patients are randomly assigned to one of the firms and remain with that firm from then on for their general medical care, inpatient and outpatient. Designed in part to provide an educational focus for general internal medicine and for continuity of care, they are an

From the Departments of Medicine, Biometry, and Community Medicine and Epidemiology, Case Western Reserve University, and the Cleveland Metropolitan General Hospital, Cleveland, Ohio.

Reprint requests: David I. Cohen, M.D., M.Sc., Cleveland Metropolitan General Hospital, Room H-331, 3395 Scranton Road, Cleveland, OH 44109.

This study was supported in part by the Cleveland Foundation, Primary Care Funds, the State of Ohio, and the Board of Regents.

0025-7079/82/0300/0286/$00.85 © J. B. Lippincott Co.

extraordinary vehicle for ongoing health services research.[3]

Study Design

We hypothesized that feeding back information about the costs of laboratory and x-ray tests to physicians would lead to reduced use of such tests. When the feedback was discontinued, there would be a decay of the effect and test ordering behavior would rise again, but not as high as its original level.

Two firms were randomly chosen to receive only cost information about lab tests, and the other two firms received cost information only about x-ray tests. Information about actual use of x-rays and lab tests was obtained from all four firms. The x-ray firms served as controls for the lab test firms and vice versa. After much thought, this study design was chosen to allow for the occurrence of a similar Hawthorne effect in all four teams. The hypothesized effects and study design are diagrammatically presented in Table 1, which shows the initial high level of testing (H) during the pre-experimental control period (April 1980). During the experimental period (May 1980), testing would fall to a low level (L) only for the feedback tests (x-rays in firms A and D, lab tests in firms B and C). In the postexperimental follow-up period (June 1–22, 1980), the effect would decay, resulting in a moderate or intermediate level of testing (M).

Literature Review

Rapid increases in expenditures for health care have focused attention on the expanding role of technology as a factor in medical inflation.[4] The diagnostic laboratories and diagnostic radiology are particularly conspicuous consumers of the health care dollar.

Many investigators have found that diagnostic test and x-ray ordering is not optimal.[5–7] Goldberg and Abbott[8] found that even if the physician does order the appropriate test, he or she may not make an appropriate response in terms of additional tests or therapies. Schroeder and O'Leary[9] report that ordering of lab tests and x-rays is higher in university-affiliated hospitals than in community hospitals. In one of the few studies to employ even a rudimentary check on quality of care (by observing mortality rates), Martin et al.[10] found that for coronary care patients in one community hospital, lab usage increased geometrically over 30 years, while mortality in severity-adjusted cases remained stable.

Increasing costs are only part of the burden of diagnostic overkill. Even noninvasive procedures carry health risks for the patient. Schimmel[11] found that almost 3 per cent of the patients in a university medical service suffered ill effects in response to diagnostic tests.

It is not yet clear what factors affect physician ordering patterns. Schroeder et al.[12] report that the quality of the interns on medical service (as ranked by faculty) has no bearing on the quantity of tests ordered. Freeborn[13] was able to correlate lab usage among experienced clinicians to the place

TABLE 1. Hypothesized Results

Team (Firm)		Control Pre-experimental Period (April 1980)	Experimental Period (May 1980)	Follow-up Period (June 1980)
A	L	H	H	H
	(X)	H	L	M
B	(L)	H	L	M
	X	H	H	H
C	(L)	H	L	M
	X	H	H	H
D	L	H	H	H
	(X)	H	L	M

H = High volume of testing, L = Low volume of testing, M = Medium volume of testing; A, B, C, D represent four firms (teams of physicians); L = Volume of lab tests, control conditions; (L) = Volume of lab tests, experimental conditions; X = Volume of x-ray tests, control conditions; (X) = Volume of x-ray tests, experimental conditions.

of their training, age and board certification. He particularly noted that more junior members of the staff tended to follow the example set by the leaders on their service.

Dixon and Laszlo[14] found that a substantial fraction of lab tests ordered had no effect on the course of patient care. In a small sample, they placed a limit on the number of tests allowed each intern. The percentage of tests found to be significant to care jumped from 5 to 23, while total usage fell by 25 per cent. No comment was made on the effect of usage on quality of care.

Many proposals have been put forth to reduce the physician's use of unneeded lab tests and diagnostic x-rays. Many of these center on greater educational effort.[15–17] Griner[18] has had some long-term success with an educational program, but he also used other techniques such as the cost feedback examined in the present study. It has also been reported[16] that the effect of educational programs may decay over the long run.

Many writers have suggested using financial incentives to induce house staff to lower the use of diagnostic facilities. Buck and White[19] report some success in reducing the use of a variety of medical services by more stringent claims review procedures.

A particularly promising approach has been to review physician behavior and feedback information (either privately[20] or publicly[21]) concerning the value of the procedures ordered. Martin et al.[22] found chart review to be superior to financial incentives in eliciting a reduction in test usage by medical residents. It appears that other pressures on the house officer (to be thorough, not to miss a diagnosis, to please the more senior physicians, etc.) are more powerful than financial ones.[5]

It is probably true that most physicians are aware of the importance of keeping hospital costs down. However, it is rare for house officers to know the actual costs involved in even the most simple and routine determinations.[15,23–24,25] Schroeder et al.[26] found that distributing a cost audit to the physician reduced lab test costs by 29.2 per cent. A more recent report[27] showed similar findings, although no attempt was made to correct for the confounding element of length of stay.

Methods

All inpatient laboratory tests, including hematology and chemistry, were recorded for all inpatients. X-ray tests included only chest x-rays, CT scans, ultrasound and liver scans. Charges were based on the hospital price list. Five days a week each firm was given its own lab tests of the preceding day by patient name, type of test ordered and total cost. Monday's report included the weekend's tests. SMA6, 12 and 18 test sets were counted as one test, not as six, 12 and 18 tests; however, their prices differed. These data sheets were handed to the physician.

X-ray tests were compiled weekly and reported by patient name, date, type of test and cost. Each firm's usage data were returned once a week and posted. No comparative data about other firms were posted. At the beginning of the experimental period a brief explanation was given to the senior resident to the effect that test usage was being reported to the firms and test usage was going to be analyzed to see if any changes occurred. During the pre-experimental period test usage was recorded unobtrusively. This recording continued through all three periods of the study. Basic demographic characteristics of the patients were also collected.

Costs and Charges

As in all of the financing of health care, the costs and charges issue is a tangled one. Because much of the cost of lab testing is fixed in the short run, a large reduction in test usage will not be reflected by a large reduction in hospital costs. The staff who

draw blood will not be laid off, and the laboratory equipment will remain in place. Long-run changes in testing will very likely result in cost savings to the hospital. However, when lab tests are charged to patients and third parties, then they become clear costs to these payors and immediate savings will occur to them. Test charges have been used here.

Patients in the firms at the start of the control period and remaining in the firms at the end of the follow-up period were excluded. This left 909 patients who entered the four firms as inpatients. Of these patients, 398 overlapped the control-experimental and experimental and follow-up periods, and were excluded. The remaining 511 patients whose stay was completely within one of the three periods were chosen for primary analysis. Of this group 52 per cent were female, the mean age was 54.5 years, and 66 per cent were white. No significant differences in these characteristics were found between firms or during the three periods of the study, suggesting adequate randomization of patients.*

Findings

Table 2 shows mean values for test usage by the 511 patients in the study. Examination of correlations associated with test usage showed the following results for these 511 patients. The number of first day lab tests is positively correlated with all lab tests ($r = 0.43$, $p < 0.01$), and with total lab charges ($r = 0.43$, $p < 0.01$). Number of lab tests is very strongly related to lab test charges: +0.920 for first day tests and +0.971 for all tests. The length of patient stay is a good predictor of test usage, +0.514 for number of lab tests, +0.483 lab test charges and +0.376 x-ray charges. There is significant but less association with CT use, +0.122 ($p < 0.05$) and ultrasound, + 0.266; length of stay is not correlated with first day lab testing. Our subjective severity-of-illness score was not associated with length of stay, but it was associated with more lab testing (+0.123, $p < 0.05$) for the first day testing, and all tests, +0.157, but not with x-ray use. Higher lab test charges were associated with more x-ray charges (+0.427). Different types of x-ray testing were not correlated with each other. The patient's insurance status had no effect on the use of tests or x-rays.

As predicted, there were no significant differences between firms during the control period. As predicted, the firms that received lab test data reduced their test usage during the experimental period ($p < 0.10$), and it returned to a similar level after the data feedback was removed in the follow-up period.

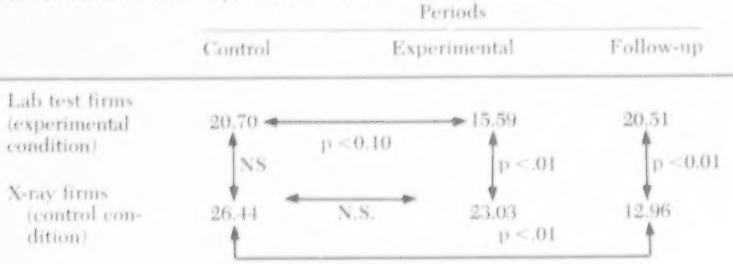
As predicted, x-ray firms did not change their use of lab tests from control to experimental conditions. The initial volume of testing in the x-ray and lab firms was not significantly different during the control period, but became so during the experi-

TABLE 2. Mean Values and Standard Deviations for Test Usage (n = 511)

	Mean	SD
Number of lab tests per admission	20.39	22.79
Lab test charges per admission	$214.81	257.97
X-ray charges per admission	$ 49.80	63.50
Number of lab tests on first day	6.28	5.11
Lab test charges for first day	$ 64.83	60.41

* Patients admitted to the four firms during the study:
There were 302 new patients without overlap between study periods and 209 admissions without overlap for patients who had been seen by the firm physicians prior to this admission, for a total of 511.
There were 398 patients overlapping study periods.
Total number of patients admitted was 909.

FIG. 1. Mean number of lab tests per admission during the control, experimental and follow-up periods for two firms with lab test information only, and two firms with x-ray test information only (total n = 511).



mental condition ($p < 0.01$). So far, these findings are consistent with our hypothesis.

What was unexpected was the abrupt and significant drop in the use of lab tests by the x-ray firms in the follow-up period. We had predicted no change, yet it fell to the lowest usage of all—half the volume of the control period. These findings are summarized in Figure 1.

Figure 2 shows the same analysis using laboratory test costs. The differences were not significant, but the pattern is exactly the same as for the number of tests shown in Figure 1. Test usage per patient day showed a similar pattern.

Although the differences in x-ray usage were not significant, there was a decline in x-ray usage during the experimental conditions. This slight decline continued in the x-ray firms but was reversed in the lab test firms. These results are summarized in Table 3.

Unlike the lab test data, which were reported back to the physicians five times a week and usually within 24 hours, the x-ray data were reported back a week later. The lack of immediate feedback to the physicians could explain the smaller effect.

Similar analysis was performed for the smaller group of 302 new arrivals, which showed a similar pattern, and similar results occurred.

Discussion

After the study, physician team (firm) leaders were interviewed concerning their perception of the study. In two firms, which happened to be x-ray firms, firm leaders took an active interest in the study. In one case the full-time chief expressed his interest through informal discussions with the other physicians on the team. In another firm the chief resident required the junior physician with the highest costs to buy the weekly box of donuts for the Friday physician firm conference, a small, symbolic, but visible gesture.

The two passive leaders were in the lab test firms. One of them said that he did not feel it was his role to involve himself in his team's test-ordering behavior. He further said that there is no room for variation in

FIG. 2. Mean laboratory test charges per admission during control, experimental and follow-up periods for two firms with lab test information only and two firms with x-ray test information only (total n = 511).

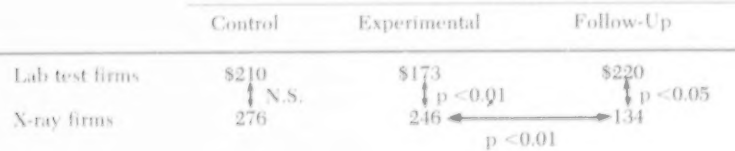


TABLE 3. Mean X-ray Charges per Admission during Control, Experimental and Follow-up Conditions for Two Firms with Laboratory Test Information Only and Two Firms with X-ray Test Information Only (Total n = 511)

	Periods		
	Control	Experimental	Follow-up
Lab test firms (control condition)	$55.34	$40.12	$49.35
X-ray firms (experimental condition)	57.37	48.21	45.07

Differences were not significant.

test ordering. A specific type of patient had to have specific tests. Therefore, information on test use and costs could have no effect on physician test-ordering behavior.

Although our findings and other studies show the contrary, his view appears to explain what happened. Cost information must, like all data, be perceived by someone whose mind is prepard to use it, and in the case of a team leader, one who is prepared to convey that message to the team. In the absence of this preparation, information will fail to be useful.

It appears that reductions in laboratory and x-ray tests occurred in the x-ray firms, even though the study design did not provide feedback on laboratory tests and the x-ray feedback was delayed. The greatest change occurred after all feedback had stopped during the follow-up period.

Our findings are quite similar to those of Martin et al.[22] They had three groups of physicians in their study: a control group; a group given a money incentive, $150 if test use dropped by 20 per cent; and a third group with a chart review of 25 per cent of the physicians' patients by a preceptor in weekly discussion groups focusing on tests the preceptor thought to be unnecessary.

We had initially thought of using a money incentive, but our human subjects committee thought it would be unethical to reward physicians not to test. Martin et al.[22] found no change in the control and money reward groups, but a sharp drop in the record review and discussion group, a drop that continued into the follow-up period during which the lowest test usage was found.

Rewards without a "prepared mind" may have led to the lack of change in their money incentive group, as was the case in our lab test groups.

The policy implications seem clear. Simple cost feedback mechanisms will not by themselves assure reduction in test usage. It requires effort to prepare physicians to use these data, an effort that is likely to be cost beneficial because of the large savings that can be achieved by lowering laboratory test charges.

References

1. Gibson RM. National health expenditures 1978. Health Care Finan Rev 1979;1:1.

2. Carels EJ, et al. The physician and cost control. Cambridge, Mass.: Oelgeschlager, Gunn & Hain, 1980;24.

3. Waggoner D, et al. A "firm" system for graduate training in general internal medicine. J Med Educ 1979;54:556.

4. Fineberg HV. Clinical chemistries: the high cost of low-cost tests in medical technology: the culprit behind health care costs? In: Altman SH, Blandon R, eds. Washington, D.C., U.S. Government Printing Office, 1979.

5. Griner PF, Liptzin B. Use of the laboratory in a teaching hospital: implications for patient care, education and costs. Ann Intern Med 1971;75:157.

6. Busse HP. The use and abuse of lab tests. Med Clin North Am 1969;53:223.

7. Abrams HL. The "overutilization" of x-rays. N Engl J Med 1979;300:1213.

8. Goldberg GA, Abbott JA. Explicit criteria for use of lab tests (Letter). Ann Intern Med 1974;81:857.

9. Schroeder SA, O'Leary DS. Differences in lab use and length of stay between university and community hospitals. J Med Educ 1977;52:418.

10. Martin SP, et al. Inputs into coronary care during thirty years: a cost-effectiveness study. Ann Intern Med 1974;81:289.

11. Schimmel EM. The hazards of hospitalization. Ann Intern Med 1964;60:100.

12. Schroeder SA, Schliftman A, Piemme TE. Variation among physicians in use of laboratory tests: relation to quality of care. Med Care 1974;12:709.

13. Freeborn DJ. Determinants of medical care utilization: physicians use of lab services. Am J Public Health 1972;62:846.

14. Dixon RH, Laszlo J. Utilization of clinical chemistry services by medical house staff. Arch Intern Med 1974;134:1064.

15. Hurst JW. Teaching cost-containment to medical house staff. J Med Assoc Ga 1978;67:889.

16. Eisenberg JM. An educational program to modify lab use by house staff. J Med Educ 1977;52:578.

17. Benson ES. Clinical lab utilization and cost of health care. Postgrad Med 1978;63:20.

18. Griner PF. Use of lab tests in a teaching hospital: long-term trends. Ann Intern Med 1979;90:243.

19. Buck CR, White KL. Peer review: impact of a system based on billing claims. N Engl J Med 1974;291:877.

20. Eisenberg JM, et al. Computer based audit to detect and correct overutilization of laboratory tests. Med Care 1977;15:915.

21. Dyck FJ. Effect of surveillance on the number of hysterectomies in Saskatchawan. N Engl J Med 1977;296:1326.

22. Martin AR, et al. A trial of two strategies to modify the test ordering behavior of medical residents. N Engl J Med 1980;303:1330.

23. Wenberg JE. Change in tonsillectomy rates. Pediatrics 1977;59:821.

24. Nagurney JT, et al. Physician awareness of economic factors in clinical decision-making. Med Care 1979;17:727.

25. Robertson WO. Costs of diagnostic tests: estimates by health professionals. Med Care 1980;18:556.

26. Schroeder SA, et al. Use of lab tests and pharmaceuticals: variation among physicians and effect of audit. JAMA 1973;225:969.

27. Henderson D, et al. Hospital cost containment: a little knowledge helps (Abstract). Clin Res 1979;27:279a.

MEDICAL CARE
March 1982, Vol. XX, No. 3

The Effects of Price Information on Physicians' Test-Ordering Behavior

Ordering of Diagnostic Tests

K. MICHAEL CUMMINGS, PH.D., M.P.H., KENNETH B. FRISOF, M.D., MICHAEL J. LONG, PH.D., AND GEORGE HRYNKIEWICH, B.A.

This research evaluated the effects of providing physicians with information about the prices of diagnostic tests on their subsequent test-ordering behavior. The study population consisted of 36 second- and third-year residents and 23 clinical faculty in three family practice centers affiliated with the Department of Family Medicine at Wayne State University School of Medicine, Detroit, Michigan. Study participants were asked to review four case studies, each describing a patient with ambiguous symptoms, and to indicate on an attached test order form the tests they would order for each patient. Subjects were randomly assigned either to a group that received test order forms on which the prices of diagnostic tests were printed (price-information group) or to a group that received test order forms with no prices indicated (control group). The study results show that for each of the four cases, the average number of diagnostic tests ordered was significantly lower in the price-information group than in the control group. Our findings also show an average reduction in the cumulative cost of tests ordered per patient of 31.1 per cent related to the provision of price information. The feasibility of regularly providing physicians with price information is discussed and reviewed in light of other approaches that have been developed to modify physician behavior in ordering diagnositc tests.

THE INCREASED USE of hospital and out-of-hospital laboratory testing has contributed substantially to the rise in health care expenditures in the past decade.[1,2] Between 1972 and 1977, expenditures for hospital laboratory tests nearly doubled, from $2.2 billion to over $4 billion.[1] Laboratory testing and radiologic services alone now account for up to 25 per cent of total patient costs at some hospitals.[3,4] The cost of laboratory work both in physicians' offices and in independent laboratories amounted to over $4.2 billion in 1978. Between 1972 and 1977, even though per capita visits to physicians decreased from 5.0 to 4.8, out-of-hospital laboratory tests (excluding x-rays and radiation therapy) increased from 850 tests per 1,000 physician visits to 1,510 tests per 1,000 physician visits.[1]

While much of the increase in the use of diagnostic testing is medically justifiable, there is evidence suggesting that part of the increase in test ordering is unnecessary. In a study carried out at the University of Rochester, Griner and Liptzin[3] examined the relationship between the number of laboratory tests ordered for patients hospitalized with diagnoses of diabetic ketoacidosis and pulmonary edema and

From the Department of Cancer Control and Epidemiology, Roswell Park Memorial Institute, Buffalo, New York; the Department of Family Medicine and Community Medicine, Wayne State University, Detroit, Michigan; and the Cost Containment Department, Blue Cross–Blue Shield of Michigan.

This research was supported by NIH grant number 05384-16.

Requests for reprints: K. Michael Cummings, Ph.D., M.P.H., Research Scientist, Department of Cancer Control and Epidemiology, Roswell Park Memorial Institute, 666 Elm Street, Buffalo, NY 14263.

0025-7079/82/0300/0293/$00.95 © J. B. Lippincott Co.

length of stay over a three-year period. Their findings showed a 27 per cent increase in the number of laboratory tests ordered but no corresponding decline in length of stay. Other studies using various criteria to determine the appropriateness of laboratory test ordering have reported substantial over-ordering of such tests as calcium determinations, barium enemas and upper gastrointestinal series.[5–7] In addition, several studies have found that test findings are often ignored by physicians, implying that the tests were not needed to begin with.[3,8,9]

One factor that has been suggested as contributing to the unnecessary ordering of laboratory tests is the physicians' lack of knowledge of the price of tests they order. A number of studies have reported that physicians are generally unaware of the price of commonly ordered diagnostic tests.[10–15] Moreover, these studies have shown a consistent tendency for physicians to underestimate the prices of diagnostic tests.

The origin of this lack of knowledge of prices may very well lie in the fact that health professionals in general are opposed to any consideration of cost in the provision of quality medical care. As Donabedian[16] points out, there is general acceptance of the idea that everything that might contribute positively to the treatment of the patient must be done. However, Donabedian further indicates that while improvements in care will almost certainly result from the increased expenditures, the marginal gain from each new expenditure is likely to diminish. For example, initial studies of cardiac enzymes can yield useful clinical information for diagnosis of myocardial infarction. However, once the diagnosis is made and treatment is initiated, daily cardiac enzyme studies provide little additional useful clinical information.

In this article we report findings from an experimental study designed to evaluate the effects of providing physicians with information about the prices of diagnostic tests on their subsequent test-ordering behavior.

Methods

Study Population

The study population consisted of 36 second- and third-year residents and 23 clinical faculty in three family practice centers affiliated with the Family Medicine Department at Wayne State University School of Medicine in Detroit, Michigan.

Intervention

Study participants were asked to review four case studies, each describing a patient's symptoms and pertinent findings from physical examinations, and to indicate on an attached order form the tests they would order for each patient. A standard test order form was used for all four patients. The cases were constructed to describe common problems seen in adult primary care at a level of ambiguity frequently seen in clinical practice.

Case studies 1 and 2 described patients seen in an ambulatory setting. The first case described a 41-year-old man under marital and job stress who complained of fatigue, weight loss and a variety of mild somatic symptoms in the respiratory, gastrointestinal and urinary tracts. Possible diagnostic considerations included depression, gastritis and alcoholism. The second case described a 56-year-old man with a two-week history of low back pain starting after unusually strenuous activity. The patient's physical examination was most consistent with low back strain, but other causes of the pain were possible.

Case studies 3 and 4 described patients who were admitted to the hospital. The third case described a 57-year-old woman who was an insulin-dependent diabetic with a two-day history of fever, chills and a productive cough. Probable diagnoses included pneumonia and uncontrolled diabetes, among others. The fourth case described a 78-year-old woman with congestive heart failure controlled with digoxin

and furosemide who ran out of the diuretic and presented with increasing dyspnea and forgetfulness. Diagnostic considerations included medication noncompliance and acute myocardial infarction.

The four case studies were reviewed by three board-certified internists whose diagnostic impressions matched the intended diagnoses for each case.

Within each family practice center, subjects were randomly assigned either to a group that received test order forms on which the prices of diagnostic tests were printed (price-information group) or to a group that received test order forms with no prices indicated (control group). A stratified random assignment procedure was used to insure a proportionate distribution of residents and faculty in each group. Because prices of diagnostic tests differed in each setting, separate price lists were developed for each family practice center.

Data Collection

One week before data collection began, residents and clinical faculty were sent a letter briefly describing the purpose of the study and urging their participation. Within the next week, subjects received in their office mailboxes the four case studies and test order forms, along with a cover letter describing how to complete the forms. Subjects were instructed to return the completed test order forms to the resident secretary within two weeks. Subjects not returning the materials after two weeks were sent a second set of materials and a letter encouraging their participation. Additional follow-up was carried out by sending daily reminder cards to subjects who had not returned the study materials. Of the 67 subjects contacted, 59 participated in the study (88 per cent response rate).

Dependent Variables

Three variables were analyzed for difference between the price-information group and the control group: 1) a count of the number of diagnostic tests ordered for each case; 2) a count of the number of tests ordered within different categories of tests (11 categories of diagnostic tests were constructed including hematology, urinalysis, chemistry, immunology, radioimmunoassay, microbiology, radiology-plain, radiology-contrast, other radiology, nuclear medicine and other); and 3) the cumulative cost of diagnostic tests selected for each case.

Analysis Plan

Two-sample t-tests were used to evaluate differences in the number of diagnostic tests ordered and the cumulative costs of ordered tests between the price-information group and the control group. Analyses were carried out separately on each case and combined across cases. An F-test was computed to evaluate the homogeneity of variances between the experimental groups. In several instances, the F-test indicated that the variances were unequal. To correct this problem the dependent variables were transformed using either a square root transformation or a logarithmic transformation. The square root transformation was applied to the variables representing the number of tests ordered, and the logarithmic transformation was applied to the variables representing cumulative cost of tests ordered.* After the variables were transformed, there were no statistically significant differences at the $p < 0.05$ level in the variances of the two groups. A nonparametric procedure, the Mann-Whitney U-test, was also used to evaluate differences in the dependent variables between the price-information group and the control group. In no case did the findings from either the nonparametric

* When the variates to be transformed include zeros, a technical problem arises, since the logarithmic transformation of zero is negative infinity. To correct this problem, the logarithmic transformation was based on the variable value plus 1.

procedure or from the two-sample t-tests using the transformed variables change the interpretation of the results obtained using the untransformed variables.† The results presented in this article are based on the untransformed data.

Results

Figure 1 shows the average number of diagnostic tests ordered for each case by experimental group. For each case, the average number of diagnostic tests ordered is lower in the price-information group than in the control group. Comparison of the number of tests ordered combining all four cases indicates that on the average, physicians in the price-information group ordered 30.8 per cent fewer tests per patient than did physicians in the control group ($p < 0.01$).

Comparison of the average number of tests ordered within 11 different categories of diagnostic test procedures indicates that in all but two instances physicians in the price-information group ordered fewer tests in each category of diagnostic test procedure than did physicians in the control group. For each of the four cases, the largest difference in number of tests ordered is in the category of blood chemistries.

To evaluate whether price information reduces overutilization of unnecessary tests or inadvertently decreases ordering of necessary tests, the price-information group and the control group are compared with regard to the ordering of necessary tests. Three board-certified family practice physicians reviewed each case study to identify the minimum necessary tests that should be ordered for each patient. The necessary tests for case 1 included a complete blood count or hemoglobin, and urine analysis. For case 2, no tests were identified as being absolutely necessary. For case 3, the necessary test identified included a complete blood count with differential, urine analysis, electrolytes, glu-

† The results from the two-sample t-tests using the transformed variables and nonparametric tests are available upon request from the author.

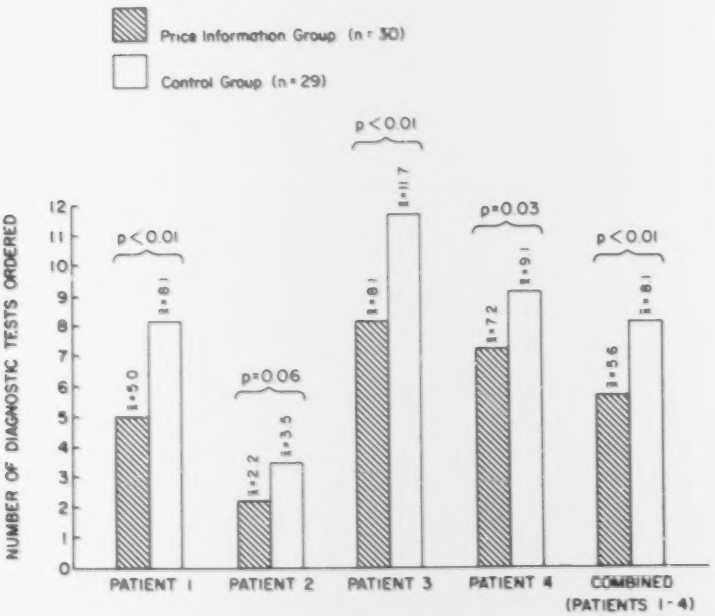


FIG. 1. Average number of diagnostic tests ordered per patient by experimental group.

cose and chest x-ray. The necessary tests for case 4 included electrolytes, chest x-ray and EKG. Table 1 shows the percentage of physicians ordering the minimum necessary tests for cases 1, 3 and 4 by experimental group. As shown in Table 1, there is little difference in the ordering of necessary tests between physicians in the price-information group and the control group.

The average cost of diagnostic tests ordered for each case by experimental group is shown in Figure 2. For each case, the average cost of tests ordered was lower in the price-information group than in the control group. The difference in the average cost of tests selected was statistically significant at the $p < 0.05$ level for case 1 and case 3, but not for case 2 ($p = 0.09$) and case 4 ($p = 0.09$). Combining the data for all four cases, there is an average reduction in the cumulative cost of tests ordered per patient of 31.1 per cent related to the provision of price information ($p < 0.01$).

The average number of diagnostic tests ordered by clinical faculty and residents is presented in Table 2. As shown in Table 2, in each of the four cases residents ordered more tests than clinical faculty, although none of the differences observed was statistically significant at the $p < 0.05$ level.

The provision of price information affected the test-ordering behavior of residents differently than it did that of clinical faculty. As shown in Figure 3, residents in the price-information group ordered significantly fewer diagnostic tests compared with residents in the control group for case 1, case 2 and case 3, but not for case 4. As indicated in Figure 4, clinical faculty in the price-information group ordered significantly fewer diagnostic tests than did their counterparts in the control group for cases 3 and 4, but not for cases 1 and 2.

Discussion

The findings from this study indicate that providing physicians with information about the prices of diagnostic tests can significantly alter their test-ordering behavior

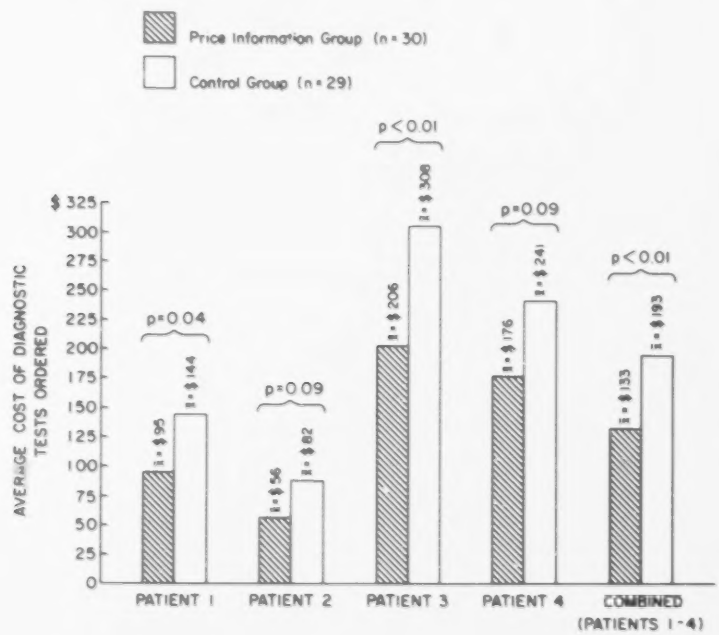


FIG. 2. Average cumulative cost of diagnostic tests ordered per patient by experimental group.

TABLE 1. Percentage of Physicians Ordering Necessary Tests Per Patient by Experimental Group

	Price Information n = 30	Control n = 29
Patient 1	80%	86%
Patient 2	53%	55%
Patient 3	53%	48%

in the direction of ordering fewer tests. Moreover, the provision of price information resulted in a consistent reduction in the number of tests ordered across several different categories of diagnostic tests (although the only consistent, statistically significant reduction in tests ordered was in the category of blood chemistries). Our findings also show a substantial cost savings related to the provision of price information. The cumulative cost of tests ordered per patient was 31.1 per cent lower in the price-information group than in the control group. However, the strong positive correlations found (product-moment correlations ranged from 0.77 to 0.90 for the four cases) between the number of tests ordered and the cumulative cost of tests ordered suggest that most of the cost savings observed was the result of a reduction in the number of tests ordered rather than of selection of less expensive alternative tests.

The differential effects of price information on the test-ordering behavior of residents and clinical faculty are not easily explained. It appears that for residents, the severity or urgency of the patient's condition is an important factor. Price information was more likely to influence the test-ordering behavior of residents for the less acutely ill patients (cases 1 and 2). It is possible that residents are more amenable to the effects of price information when there is less risk associated with not having complete information about the patient's condition. In contrast, for clinical faculty price information had little effect on test-ordering behavior for the commonly seen ambulatory problems (cases 1 and 2). Physicians who have been in practice for some time often have established relatively consistent patterns for diagnosing common conditions and thus are less susceptible to attempts to modify their test-ordering behavior relative to these conditions. A separate explanation is that clinical faculty are more knowledgeable about the prices of commonly ordered diagnostic tests than are residents, and are therefore not as likely to be affected by the provision of price information.

The generalizability of this study's results is limited, both with respect to the sample and by the use of fictitious case studies. The physicians in this study were all family practitioners associated with a

TABLE 2. Average Number of Diagnostic Tests Ordered by Clinical Faculty and Residents

	Clinical Faculty n = 23 x	Residents n = 36 x	Difference	p-value
Patient 1	6.3	6.7	−0.4	p = 0.75
Patient 2	2.5	3.1	−0.6	p = 0.39
Patient 3	9.1	10.4	−1.3	p = 0.21
Patient 4	7.6	8.5	−0.9	p = 0.31
Combined (patients 1–4)	6.3	7.2	−0.9	p = 0.29

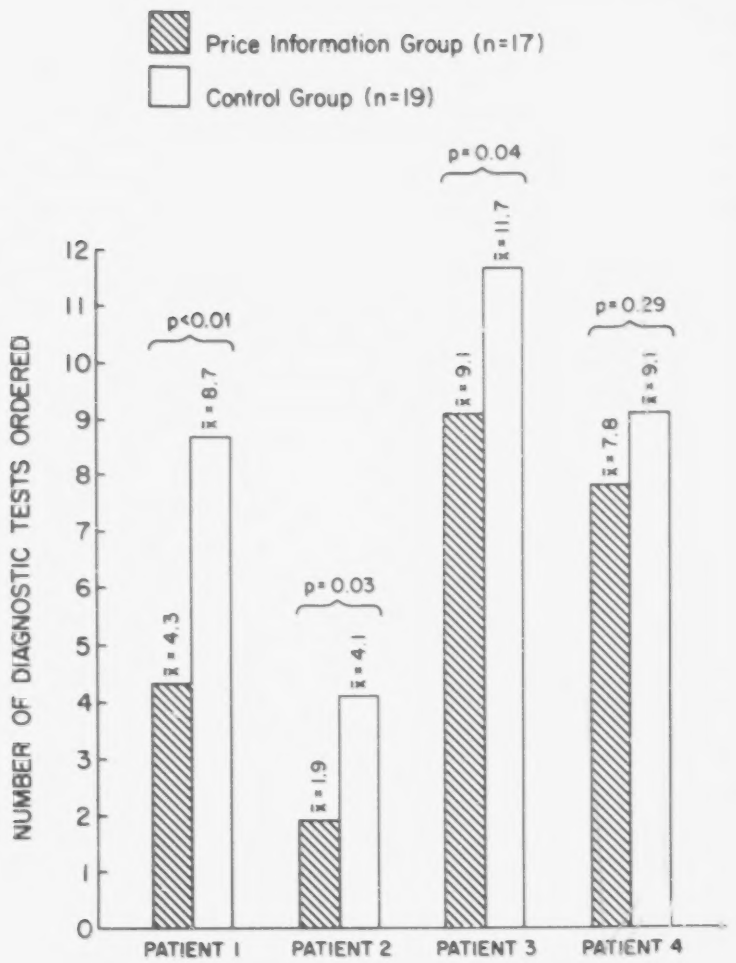


FIG. 3. Average number of diagnostic tests ordered per patient by residents and experimental group.

medical residency training program. It is not clear whether the results obtained with this group of physicians would transfer to other primary care specialty groups or to physicians working in nonacademic settings. It is also not clear whether the results obtained using case study examples accurately reflect physician behavior with their patients. There is some evidence suggesting that the use of written case simulations does provide a good indication of how physicians respond to actual patient problems.[24,25] A further limitation of the case study format used in this research is that it did not allow an opportunity to observe repeat test ordering (i.e., additional test ordering after the results of initial tests are obtained). In a study of the test-ordering behavior of first-year medical residents, Martin et al.[19] reported that repeat testing accounted for about 75 per cent of all laboratory tests ordered for patients in a hospital setting over a one-year period.

Future research should concentrate on evaluating the impact of providing physicians with price information on the number and types of diagnostic tests ordered for actual patient problems and on

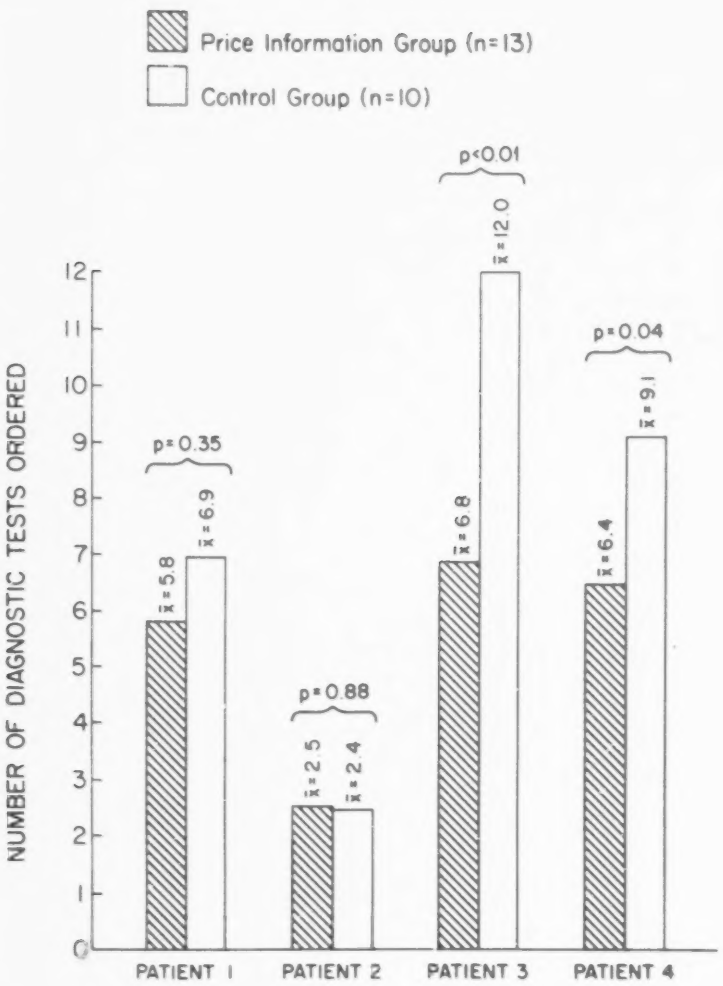


FIG. 4. Average number of diagnostic tests ordered per patient by clinical faculty and experimental group.

the effects of price information in different patient care settings (i.e., hospital, outpatient clinics, private practice, health maintenance organizations). Additional research is also needed to determine which physicians are susceptible to the effects of price information. For example, Schroeder et al.[14] found that a cost audit with feedback to physicians was most effective in reducing the number of tests ordered by "high use" physicians. In our study, the lower variability in the number of tests ordered by physicians in the price information group compared with the control group may reflect a tendency on the part of "high use" physicians to restrain their test ordering when exposed to price information.

Several approaches to modification of physician behavior in ordering diagnostic tests have been reported in the literature. These methods have included setting limits on the number of tests ordered in a day,[8,17] offering incentives to physicians for reduced test ordering,[19] educating physicians about the appropriate ordering of tests,[5,21,22] auditing physician test-ordering practices with feedback to physicians about appropriate test ordering[18–20] and providing daily feedback of patients' bills to the physicians.[14,23] Evaluations of

the effectiveness of these different approaches in modifying physicians' test-ordering behavior report reductions in test ordering ranging from 0 to 36 per cent. Overall, auditing of physicians' test-ordering behavior along with feedback concerning appropriate test ordering appears to be the most consistently effective strategy for modifying physicians' test-ordering behavior. Unfortunately, while chart reviews appear to be effective, they are expensive to implement (for example, Martin et al.[19] estimated that it would cost $15,000 per year to conduct six chart review sessions with first-year medical residents on the medical service of a large hospital) and may be resisted if physicians feel that their autonomy and responsibility for patient care are being threatened.

The results of this study suggest that providing physicians with information about the prices of diagnostic tests (before tests are ordered) is an effective strategy for reducing the number of tests that are ordered. While this approach does not deal directly with the problem of the inappropriate ordering of tests, our findings indicate that price information does not affect ordering of necessary tests. In comparison with other approaches for modifying the test-ordering behavior of physicians, providing information on the prices of diagnostic tests is a relatively simple and inexpensive intervention to implement.

References

1. Gibson RM. National health expenditures. Health Care Fin Rev 1979;1:1.

2. Relman AS. The allocation of medical resources by physicians. J Med Educ 1980;55:99.

3. Griner PF, Liptzin B. Use of the laboratory in a teaching hospital: the implications for patient care, education, and hospital costs. Ann Intern Med 1971;75:157.

4. Schroeder SA, O'Leary PS. Differences in laboratory use and length of stay between university and community hospitals. J Med Educ 1977;52:418.

5. Eisenberg JM, Williams SV, Gorner L, et al. Computer-based audit to detect and correct overutilization of laboratory tests. Med Care 1977;15:915.

6. MacEwan DW, Kavanagh S, Chow P, Tishler JM. Manitoba barium enema efficacy study. Radiology 1978;126:39.

7. Morton KI, Sox HC, Wasson J, Duisenberg LE. The clinical value of the upper gastrointestinal tract roentgenogram series. Arch Intern Med 1980;140:191.

8. Dixon RH, Laszlo J. Utilization of clinical chemistry services by medical housestaff. Arch Intern Med 1974;134:1064.

9. Williamson JW, Alexander M, Miller GE. Continuing education and patient care research: physician response to screening test results. JAMA 1967;201:118.

10. Dresnick SJ, Roth WI, Linn BS, et al. The physician's role in the cost containment problem. JAMA 1979;241:1606.

11. Freeman RA. Cost containment. J Med Educ 1976;51:157.

12. Kelly SP. Physician's knowledge of hospital costs. J Fam Prac 1978;6:171.

13. Robertson WO. Costs of diagnostic tests: estimates by health professionals. Med Care 1980;18:556.

14. Schroeder S, Kendeis K, Cooper J, et al. The use of laboratory tests and pharmaceuticals: variation among physicians and effect of cost audit. JAMA 1973;225:969.

15. Skipper JK, Smith G, Mulligan JL, et al. Physicians' knowledge of cost: the case of diagnostic tests. Inquiry 1976;13:194.

16. Donabedian A. Some basic issues in evaluating quality of health care. In: Issues in Evaluation of Health Care. Kansas City, Mo.: American Nurses' Association, 1976.

17. Gray G, Marion R. Utilization of a hematology laboratory in a teaching hospital. Am J Clin Pathol 1973;59:877.

18. Lyle CB, Bianchi RF, Harris JH, Wood ZB. Teaching cost containment to house officers in Charlotte Memorial Hospital. J Med Educ 1979;54:856.

19. Martin AR, Wolf MA, Thibodeau LA, et al. A trial of two strategies to modify the test-ordering behavior of medical residents. N Engl J Med 1980;303:1330.

20. Mitchell JH, Hardacre JM, Wenzel FJ, Lohrenz FN. Cholecystectomy peer review: measurement of four variables. Med Care 1975;13:409.

21. Klein L, Charache P, Johannes R, Lewis C. Effect of physician tutorials on prescribing patterns and drug cost in ambulatory patients. Clin Res 1980;28:296A.

22. Rhyne RH, Gelbach SH. Effects of educational feedback strategy on physician utilization of thyroid function panels. J Fam Prac 1979;8:1003.

23. Henderson D, D'Alessandr E, Westfall B, et al. Hospital cost containment: a little knowledge helps. Clin Res 1979;27:279A.

24. Helfer RE, Slater CE. Measuring the process of solving clinical diagnostic problems. Br J Med Educ 1971;5:48.

25. Williamson JW. Assessing clinical judgment. J Med Educ 1965;40:180.

MEDICAL CARE
March 1982, Vol. XX, No. 3

Physician Sensitivity to the Price of Diagnostic Tests:

A U.S.–Canadian Analysis

JOHN HOEY, M.D., JOHN M. EISENBERG, M.D., M.B.A., WALTER O. SPITZER, M.D., M.P.H., AND DUNCAN THOMAS, PH.D.

A questionnaire containing 11 patient management problems was completed by 495 physicians and medical students at an American and a Canadian medical school. Respondents indicated whether they would order a particular diagnostic test in each case, given five different prices for the test. Approximately 25 per cent of attending staff and a higher proportion of residents, interns and clerks responded that they would order the test depending on its price. Approximately 50 per cent of attending staff and smaller proportions of residents, interns and clerks indicated that they would not order the test even if there were no price. Respondents in Montreal were more likely than those in Philadelphia to select a price-sensitive response, the reverse of the expected tendency. Since some tests may be ordered on the basis of price, education of physicians regarding the price of diagnostic tests may alter their use of these services, but a large proportion of tests are ordered because of clinically absolute reasons, which may be insensitive to price.

DURING THE PAST TWO DECADES, the volume and cost of diagnostic investigations have sharply increased. In 1977, it was estimated that clinical laboratories consumed approximately $11 billion in health resources in the United States and that the number of laboratory tests has increased at an annual rate of approximately 15 per cent.[1] Concerns have also been raised regarding the cost and usefulness of radiologic examinations.[2,3] It has been proposed that physician education is an important factor in the use of diagnostic tests, and programs have been initiated to educate physicians about the price of diagnostic studies.[4–6]

The assumption that physicians will use laboratory tests differently as they learn the price of the tests requires a prior assumption that physicians are concerned about the cost of medical care. This implies that

From the Department of Epidemiology and Health, McGill University, Montreal, Canada; and the Section of General Medicine, Department of Medicine, and the Leonard Davis Institute of Health Economics, University of Pennsylvania, Philadelphia, Pennsylvania, U.S.A.

This work was supported in part by Health and Welfare Canada (Drs. Hoey and Thomas), the W.K. Kellogg Foundation (Dr. Spitzer) and the National Health Care Management Center at the Leonard Davis Institute of Health Economics, DHEW Grant No HS-02577 (Dr. Eisenberg).

Reprint requests: John Hoey, M.D., Department of Epidemiology and Health, McGill University, 3775 University Street, Montreal, PQ, Canada H3A 2B4.

0025-7079/82/0300/0302/$00.80 © J. B. Lippincott Co.

their use of tests is sensitive to cost. The present study was designed to test whether physicians are sensitive to the price of diagnostic tests and whether physicians differ in this sensitivity depending on their level of training and the health care system in which they practice.

The study was initially performed at McGill University in Montreal, Quebec. Since all hospital and diagnostic test costs in Quebec are fully insured, it was of interest to compare the Quebec findings with those in a fee-for-service setting where costs are born entirely or partly by a least a portion of patients, particularly for outpatient care. Therefore, the study was duplicated at the University of Pennsylvania in Philadelphia, Pennsylvania. Results from McGill and Pennsylvania are presented in this article.

Methods

Study Subjects

The study at McGill was conducted in its five hospital teaching units. At each hospital, random samples of all attending staff, interns and residents in the third residency year were drawn. The sampling fractions were varied to reflect the proportion of physicians in each sample. The sampling fractions also reflected the distribution of physicians among the various specialty departments (Medicine, Surgery, Obstetrics/Gynecology, etc.). Both salaried and nonsalaried faculty were included. Physician administrators, psychiatrists, pathologists, and those on leave of absence or sabbatical were excluded. All family medicine attending staff, interns and residents were included. In addition, a 50 per cent random sample of fourth-year students was drawn. At the Hospital of the University of Pennsylvania, a random sample of students, residents and attending staff was selected in a fashion similar to that used at McGill, except that no family medicine physicians were included.

At McGill, 353 of the 510 subjects returned usable questionnaires (response rate, 69 per cent). At the University of Pennsylvania, 142 of 257 subjects (55 per cent) returned usable questionnaires.

Questionnaire

A questionnaire* was mailed to all study subjects. The questionnaire asked subjects to respond to 11 patient management problems, each of which asked the respondent whether he or she would never order a particular diagnostic test in that clinical situation, would order the test even if it cost more than a specified amount, or would order the test depending on its price. In each case, five different prices of the test were described to determine the physician's perception regarding the relative value of the tests.

The following example is problem 1 from the questionnaire:

> A 32-year-old previously healthy male storekeeper complains of a productive cough of six days, mild fever and malaise. He has rales in the left base, temperature 101°F, pulse 100 regular. He is otherwise well, and the remainder of the physical exam is normal. Would you order a chest x-ray for this patient?
>
> a. Yes, if the chest x-ray cost less than $5.
> b. Yes, if the chest x-ray cost between $5 and $15.
> c. Yes, if the chest x-ray cost between $15 and 30.
> d. Yes, if the chest x-ray cost between $30 and $60.
> e. Yes, even if the chest x-ray cost more than $60.
> f. No, I would not order a chest x-ray.

* Available from the National Auxiliary Publications Service, c/o Microfiche Publications, 304 E. 46th Street, New York, NY 10017.

Table 1 shows how these responses were classified. In the first column we show the responses ("a"–"d") of the subjects who were "price-sensitive." Their responses indicate that they would perform the test only if test price did not exceed a stated amount. Response "e" is shown in the second column and can be termed not cost conscious, or "clinically absolute." These respondents would order the test even if it cost more than the stated amount. Response "f," shown in the third column, is also a "clinically absolute" response. For these respondents the test would not be ordered, even if it entailed no price.

Analysis

A four-way contingency table was used to describe the effects in the data, taking physician groups, cities, questions and response categories as factors. This method produces chi-square tests describing the significance of each combination of effects, considered alone and after adjustment for the other effects that are present in the data. By fitting a log-linear model including only the relevant effects, estimates of the frequency of each response category could be obtained, broken down by each factor separately and adjusted for the others.

Results

With all 11 questions, both cities and all physician groups included for analysis, the price-sensitive options were selected in 33.9 per cent of cases, and the clinically absolute responses, "yes, even if more than the specified amount" and "no, would not order test," in 23.1 per cent and 42.9 per cent, respectively. Though consistent patterns of differences were apparent between questions, cities and professional groups as described below, there remained marked disagreement among individual physicians in response to the same questions in every subgroup. Only one question produced no price-sensitive responses in some subgroups, and even for that question there remained disagreement between "always" and "never" options.

By far the largest effect was a marked variation in the responses to individual patient management problems. For example, overall, 54.5 per cent of respondents to the first question elected a price-sensitive response, and 45.5 per cent a clinically absolute response. In contrast, a patient management problem addressing the use of an electroencephalogram in a 2-year-old girl following a generalized convulsion evoked 11.9 per cent price-sensitive options. These differences among the 11 questions

TABLE 1. Responses for All Patient Management Problems by Professional Category of Respondent

Professional Category	Responses			
	Price Sensitive	Clinically Absolute: "Yes even if more than specified amount"	Clinically Absolute: "No would not order test"	Total
Clerks	37.7%	24.6%	37.7%	904
Interns	37.1	24.1	38.8	1055
Residents	33.5	21.6	44.9	1549
Staff	23.8	27.7	48.5	1763
Total	1788	1220	2263	5271

$\chi^2 = 110.88$; df = 6; $p < 0.001$.

were highly significant ($\chi_{20}^2 = 985.8$, $p < 0.00001$). However, since all physicians in both cities were asked the same questions, these differences could have had no confounding effects on comparisons between cities and physician groups.

On the other hand, differences in response rates produced quite different distributions of physicians in the two cities. In Philadelphia, the proportions of clerks, interns, residents and attending staff were 15.1 per cent, 30.6 per cent, 31.2 per cent and 23.1 per cent, whereas in Montreal they were 18.0 per cent, 15.7 per cent, 28.6 per cent and 37.7 per cent, respectively. As these differences could confound comparisons of response categories, the following results are based on the adjusted proportions obtained from a log-linear model with all significant factors included.

Table 1 shows the adjusted response frequencies in the four physician groups, summed across all patient management problems and across the two cities. These results show two important trends. First, the price-sensitive options were selected most often by the clerks and least often by the staff, with a smooth trend between the interns and residents. Second, among the clinically absolute responses, the "would not order the test" option was selected more commonly by all physician groups, but with increasing frequency among the more senior groups. These overall differences were highly significant ($\chi_6^2 = 110.88$ for partial association, $p < 0.0001$). There were also significant differences between the responses of the physician groups to different questions, but we could find no systematic pattern to such differences.

Table 2 describes the differences in responses between the two cities, summed across problems and physician groups. Montreal physicians chose the price-sensitive option significantly more frequently (35.6 per cent after adjustment) than did Philadelphia physicians (27.7 per cent, $\chi_2^2 = 17.24$, $p < 0.0002$), but the relative frequencies of the "never" clinically absolute options were similar in the two cities. Again there were significant differences between cities in the responses to the different questions, but no consistent pattern of differences could be found.

Discussion

Physicians in many countries are being asked to reduce medical care costs. Laboratory or ancillary investigations are important aspects of such costs. Hence, the mandate for physicians is essentially that they determine whether the laboratory investigation (or other procedure) will yield a benefit to the individual patient that is worth the price.

The results of this study, performed in the United and in Canada, suggest that in

TABLE 2. Responses for All Patient Management Problems in Philadelphia and Montreal

City	Responses			
	Price Sensitive	Clinically Absolute: "Yes even if more than specified amount"	Clinically Absolute: "No would not order test"	Total
Philadelphia	27.7%	31.8%	40.5%	1534
Montreal	35.6	23.3	41.2	3737
Total	1788	1220	2263	5271

$\chi^2 = 17.42$; df = 2; $p < 0.0002$.

both countries a substantial proportion of physicians are willing (at least on a questionnaire) to take price into account and to allow the price of laboratory tests to influence their clinical decisions. This tendency was more marked for clinical clerks and interns, but was also present for attending staff physicians. In 11 patient management problems, an average of almost 25 per cent of attending staff indicated that they would order a laboratory test only if the price for the test were below a specified dollar amount; for clinical clerks and interns, the average so indicating was 37 per cent, with residents occupying an intermediate position. Of note, attending physicians more often indicated that they would not order the test at any price, thus implying that the test would not be clinically useful.

It is possible that attending physicians' reluctance to order tests was due to their clinical experience and to their placing less reliance on laboratory tests than do younger physicians. It has been shown elsewhere that older physicians order fewer laboratory tests and roentgenograms than do younger physicians.[7] It is also possible that younger physicians were most price-sensitive because of recent educational efforts to teach about health care costs. However, it is of note that this study was completed before the initiation of a cost-containment educational program at the University of Pennsylvania. It may also be that a physician's decision to recommend a test as being absolutely necessary incorporated factors other than clinical ones, such as perceived patient demand, habit or defensive medicine. Alternatively, the physician may have been influenced by these considerations more when the test was neither clearly necessary nor unnecessary. In this latter case, the inference is that factors such as defensive medicine would interact with price in determining the physician's decision to order a diagnostic test.

In Quebec, all citizens are completely insured for all investigative procedures. In Philadelphia, few patients have complete insurance coverage for all investigative procedures, particularly for outpatient services. The expected tendency, for Philadelphia respondents to be more price-sensitive, was not found. In fact, Philadelphia respondents were somewhat less likely to indicate that they would order the test depending on its price and more likely to be clinically absolute. It is possible that Philadelphia physicians were more willing to order the test at a price higher than the maximum price listed because they are more accustomed to considering the charge for a service to be its price and since charges for diagnostic services may exceed their true cost. It is also possible that Philadelphia physicians, who may be accustomed to caring for inpatients with extensive third-party coverage, are not aware of the degree to which their patients' diagnostic studies are not covered, particularly when performed in an ambulatory setting. If this were the case, they would be expected to approximate the price-sensitivity of Montreal physicians, but not to fall below it. It is of note that Philadelphia physicians were less price-sensitive than Montreal physicians for diagnostic tests in the seven patient management problems that were characteristically outpatient problems, as well as for those four that were typically inpatient problems.

The results suggest that attending staff, residents, interns and clinical clerks are often willing to take price into account when ordering laboratory tests. Therefore, education of physicians regarding the cost of medical care may influence their use of diagnostic services. However, a large proportion of laboratory tests were considered to be absolutely necessary or unnecessary, regardless of price, particularly by the more senior physicians. Therefore, the potential effect of cost education may be limited. These data must be interpreted with

caution since they are based on simulated cases and since all respondents were affiliated with major teaching institutions.

Acknowledgments

The authors wish to thank Dr. A. Ryder-Cook, Ms. S. Kerrouac, Ms. S. Kenmore and Ms. D. Meltzer for their assistance in the field work.

References

1. Fineberg HV. Clinical chemistries: the high cost of low cost diagnostic tests. In: Altman SH, Blendon R, eds. Medical Technology: The Culprit Behind Health Care Costs? Proceedings of the 1977 Sun Valley Forum on National Health; 144–165. (DHEW Publ. No. (PHS) 79-3216.)

2. Abrams HL. The overutilization of x-rays. N Er g l J Med 1979;300:1213.

3. Hall FM. Overutilization of radiological examinations. Radiology 1976;120:443.

4. Skipper JK, Smith G, Mulligan JL, et al. Physicians' knowledge of costs: the case of diagnostic test. Inquiry 1976;13:194.

5. Dresnick SJ, Roth WI, Linn BS et al. The physicians' role in the cost containment problem. JAMA 1979;241:1606.

6. Nagurhey JT, Braham RL, Reader GG. Physician awareness of economic factors in clinical decision making. Med Care 1979;17:727.

7. Eisenberg JM, Nicklin D. Use of diagnostic services by physicians in community practice. Med Care 1981;19:297.

MEDICAL CARE
March 1982, Vol. XX, No. 3

Why Veterans Choose Veterans Administration Hospitalization:

A Multivariate Model

WILLIAM FRANK PAGE, PH.D*

Data from a national survey of veterans are analyzed to determine how age, income, health insurance and service-connected disability influence the veteran's choice of Veterans Administration (VA) hospitalization. Log-linear analysis is used to provide a model that expresses in numerically quantifiable terms (odds and odds ratios) the effects of these factors on the choice of VA hospitalization. Moreover, the multivariate nature of the analysis provides the opportunity for assessing the magnitude of these effects simultaneously in the presence of all other effects. The final model shows that age and health insurance act independently of each other and of all other factors. The effect of income is modified by service-connected disability status and vice versa; this interaction of two factors is probably due to VA hospital care eligibility criteria. Overall, the single most important factor affecting a veteran's choice of hospitalization is health insurance: veterans without health insurance have nearly five times greater odds of going to VA hospitals than do those with health insurance—regardless of age, income or service-connected disability status.

ACCORDING TO A National Center for Health Statistics publication,[1] in 1973 the Veterans Administration (VA) accounted for 8.2 per cent of the days of care given in general hospitals other than children's hospitals, and for 11.8 per cent of the days of care in psychiatric hospitals. These figures provide an overall measure of the nationwide magnitude of the VA hospital system.

As the veteran population ages—by 1990 more than half of the U.S. male population over the age of 65 will be veterans[2]—the pressures on the VA hospital system will grow. Furthermore, because veterans over the age of 65 without service-connected disabilities are eligible for VA hospital care by virtue of their age,† the question of whether veterans choose to receive care in VA hospitals may become a critical issue.

In this article we will study the effects of various factors that influence the choice of VA hospitals. We will use a log-linear model analysis because it enables us to express in numerical terms the size of the effects of factors while controlling for the presence of other factors, which is particularly important in analyzing a multifactor issue such as hospital utilization.

* Chief, Biometrics Division, Reports and Statistics Service, Veterans Administration Central Office, Washington, D.C.

The opinions of the author do not reflect official Veterans Administration policy and should not be so construed.

Reprint requests: Dr. William F. Page, Chief, Biometrics Division (712), Veterans Administration, 810 Vermont Ave., N.W., Washington, DC 20420.

† The ability to pay for VA hospitalization is not a determining factor for veterans without a service-connected disability if they are aged 65 years or over (see "Choice of Hospital" section for more detail).

0025-7079/82/0300/0308/$01.15 © J. B. Lippincott Co.

Despite the analytic methodology, this article is basically descriptive. We cannot expect to master the complex issue of VA hospitalization in any single analysis or report. Nevertheless, a descriptive study of this sort, providing quantified estimates of the effects of various factors in the presence of other factors, can provide a jumping-off point for future, more definitive studies.

Methodology

Data Source

In 1978, the VA conducted the first National Survey of Veterans.[3] The survey was designed as a follow-up survey to the Bureau of the Census' Current Population Survey (CPS). The CPS is a monthly nationwide survey of about 45,000 households, with the same household included in the sample for several months (4 months in, 8 months out, 4 months in again). Survey of Veterans questionnaires were mailed out to men who reported themselves as veterans in the CPS and who were in the sample for their last time—the so-called "retiring rotation group." Approximately 15,000 questionnaires were mailed out, resulting in a final sample of approximately 11,500 respondents, although this article analyzes only the data from the 1,260 veterans who indicated that they had been hospitalized in the past 12 months. More detail on the survey may be found in the VA report.[3]

Variables

Choice of the variables of interest for the multivariate analysis was made difficult by the small sample size, which made it necessary to limit the number of variables and the complexity of their coding (the more variables and categories, the sparser the data in each category). Hence, the variables in the analysis had to be limited to those with the greatest substantive interest. The group of hospitalized veterans in this analysis was categorized on the basis of five factors (all self-reported): 1) type of hospital (VA or non-VA; persons with multiple hospital episodes were included in the VA category if *any* episode was in a VA hospital); 2) health insurance‡ (yes or no); 3) age (under 45 years, 45–64, 65 and over); 4) income under $10,000, $10,000 and over); and 5) service-connected disability.§

Choice of Hospital

The issue of choice of hospital (VA or non-VA) deserves further discussion since it includes such factors as eligibility and priority for VA care. VA eligibility criteria are relatively complex[4] but may be paraphrased as follows: veterans discharged from active military service under conditions other than dishonorable are eligible; veterans needing hospitalization for service-connected disabilities have top priority; those with service-connected disabilities who need hospitalization for unrelated conditions have second priority; other veterans may be admitted for necessary hospitalization if they state under oath that they are unable to defray the cost of hospital charges elsewhere and if a bed is available; the "inability to pay" requirement does not apply to veterans 65 years of age or older or those in receipt of VA pension.

Given these eligibility and priority criteria, certain hypotheses come to mind. For example, service-connected veterans, since they have higher priority for VA hospitalization, should choose VA hospitals more often. Similarly, non-service-connected veterans under the age of 65 with higher incomes and/or health insurance should choose VA hospitals less often. We cannot say, however, that the latter group is, by definition *ineligible* for VA

‡ Including Medicare/Medicaid.

§ This factor identifies veterans who have been judged by the VA to have a disability that is service related.

hospital care, since the inability to defray hospitalization costs is a complex, case-by-case issue that may depend upon such factors as insurance deductibles, coinsurance, limitation of benefits, etc. Hence, we will analyze the data with choice of hospital as the dependent (outcome) variable, without concern for how choice of hospital might depend upon VA hospitalization eligibility and priority.

Log-Linear Model Analysis

Log-linear analysis is one of several multivariate methodologies used to analyze data with a discrete dependent (outcome) variable. These methodologies all allow one to express in numerical terms the size of the effects of various factors on choice of hospitalization and quantify those effects in a multivariate manner by assessing the impact of a particular variable, taking into account the presence of all other variables. This last point is particularly important in analyzing a complex multivariate problem such as hospital utilization, which can easily be misinterpreted and misunderstood in a univariate analysis.

There are at least two choices of methodology other than log-linear analysis with the above properties—weighted least squares and logistic multiple regression.[‖] All three methods are "asymptomatically equivalent"[5] (i.e., the differences between results grow smaller as sample sizes grow larger), and the author has chosen log-linear mostly as a personal preference. It should be pointed out, however, that logistic multiple regression will handle continuous and discrete variables, unlike log-linear analysis, which requires all variables to be discrete as does weighted least squares. One can make a substantive case for preferring age as a categorized variable, namely that the age of 65 is a "breakpoint" in defining priority for care, and although there is no substantive argument for income categorization, discrete variables are, in the author's opinion, easier to handle. Analyses with discrete variables offer the advantage of easy computation of odds ratios[8] (definitions and examples are presented in the "Results" section) and, in addition, make use of contingency table frequencies as input, so that one need not process the raw data file more than once, which, although not very important in an analysis of this size, might be a factor in analyzing large files. Availability and ease of use of computer software were also factors in the author's choice of a log-linear methodology.

The analysis was performed using the computer program BMDP3F,[9] with type of hospital treated as a dependent variable. Thus, all models contain a term that fixes the marginal distribution of the independent variables.[10] The actual cell frequencies in the program are the "raw data" in the Appendix and will allow the interested reader to replicate this analysis. # The raw data were derived from the weighted survey estimates, although the analysis treated the data as though they had come from a simple random sample.

The last point requires further clarification. Current log-linear methodology is still in developmental stages concerning the analysis of data from other than simple random sampling (the CPS is a multistage sample and so is a complex rather than simple random sample). Because this analysis is primarily descriptive in nature, we have analyzed the data as though they were from a simple random sample; this means that our p-values are only approximate, not a serious shortcoming for this descriptive analysis (a recent article on oc-

‖ When one assumes that the independent variables have a multivariate normal distribution, logistic multiple regression is equivalent to linear discriminant function analysis; this assumption may not be realistic in general and may lead to considerable bias in the linear discriminant function estimates.[6,7]

0.5 was added to each cell to eliminate zero cells, a traditional practice.

cupational changes, based on data from a supplemental survey to the CPS, takes this approach[11]).

Results

Table 1 shows the percentage of hospitalized veterans with hospital episodes in VA hospitals for the different age, service-connected disability, insurance and income categories. These tables show that older veterans, service-connected veterans, uninsured veterans and poorer veterans are more apt to use VA hospitals. Unfortunately, these simple univariate tables do not provide a complete enough picture. For example, veterans over 65 years of age are more likely to be retired and hence have lower incomes; this increases the VA hospital "market share." On the other hand, these same older veterans are more apt to have insurance (Medicare), which means a reduced VA market share. The only way to disentangle complex analytical situations like this is to perform a multivariate analysis. Log-linear model analysis allows this type of multivariate analysis, which is an important reason for its choice.

To date, no one log-linear model fitting procedure has been adopted as the standard. We have chosen an iterative procedure involving both "stepping-up" (adding variables) and "stepping-down" (deleting variables). The first step is to find a starting model, since looking at all possible models is too costly a process for this data set.

The first step of model fitting was to establish a basic idea of the complexity of the model. Following traditional techniques we fit the model with all one-variable effects[¶] (i.e., the effect of a single independent variable on the dependent variable, hospitalization), then all two-variable effects, and all three-variable effects. The results are shown in Table 2.

[¶] All models contained the term that fixed the marginal distribution of the independent variables.

TABLE 1. Percentage of Hospitalized Veterans with Hospital Episodes in VA Hospitals by Age, Income, Insurance and Service-Connected Disability

Category	Per Cent with VA Hospital Episodes	Total Number of Hospitalized Veterans in Category
Total	16.5	1260.
Age		
Under 45	11.2	373.
45–64	18.6	678.
65 and over	19.2	209.
Income		
Under $10,000	29.3	527.
Over $10,000	7.3	733.
Insurance		
Yes	10.6	1036.
No	43.6	224.
Service-connected disability		
Yes	38.4	211.
No	12.1	1049.

Table 2 shows that the three models in question all fit the data well, with none of the goodness-of-fit chi-squares being significant. In the interest of finding the simplest model, this means that we should start with the "all one-variable effects model."

In the next step, we see if all of the one-variable effects are necessary to the model. We do this by fitting four models, each of which deletes a single one-variable effect. Table 3 shows that each of the independent factors had a significant effect on hospitalization, i.e., each one-variable effect

TABLE 2. First Step in Model Fitting Lack-of-Fit Parameters

Model	X^{2*}	df	p-value
All one-variable effects	20.10	18	0.32
All two-variable effects	8.01	9	0.53
All third-variable effects	0.51	2	0.77

* Only likelihood ratio chi-squares are reported in this paper.

TABLE 3. Second Step in Model Fitting (Step-Down)

Effect Deleted From Model	Change in Chi-Square Due to Deleted Effect		
	X^2	df	p-value
Age	4.88	2	0.09
Service-connected status	46.13	1	<0.0001
Income	53.81	1	<0.0001
Insurance	68.47	1	<0.0001

was significant, although age is significant only at the p = 0.10 level.

Since it is clear that the "all one-variable effects" model fits and that no single one-variable effects can be deleted, the next step is to see if any two-variable effects should be added. In Table 4 we look at the effect of adding separately each two-variable effect to the "all one-variable effects" model.

Using the nominal p-value of 0.10, only the income and service-connected status effect is significant. This single two-variable effect is hence added to the model. The next step was to add two-variable effects in addition to income and service-connected status—none was significant. The final step was to double check the model by deleting the effects already in the model. All were significant. This left the final model with all single-variable effects plus the two-variable effect of income and service-connected status.

TABLE 4. Third Step in Model Fitting (Step-Up)

Effect Added to Model	Change in Chi-Square Due to Added Effect		
	X^2	df	p-value
Age and SC status	3.68	2	0.16
Insurance and SC status	0.82	1	0.37
Income and SC status	4.69	1	0.03
Age and insurance	0.78	2	0.68
Income and insurance	0.69	1	0.41
Age and income	2.52	2	0.28

For this model none of the residuals (standardized or Freeman-Tukey deviates) were significant, indicating that the model resembled fairly closely the original data in each cell entry. The fitted data are displayed in the Appendix. The lack-of-fit chi-square for the final model was 15.42, with 17 degrees of freedom (p = 0.57), which indicates a well-fitting model.

The final model indicates that all four independent variables had a significant impact upon choice of hospital. Age and health insurance acted independently of each other and of the remaining factors. The service-connected disability factor and the income factor were not independent, so that the effect of the one factor was dependent upon the level of the other. In the usual log-linear terminology, the model included age, insurance, service-connected status, income and the service-connected status by income interaction. We now discuss in turn the effects of each of the factors on the choice of hospital.

Age

Age had a significant effect on the choice of hospital, but it had the weakest effect of the four variables. In general, those veterans under 45 years of age chose VA hospitals less frequently than did the other age groups. Veterans aged 65 years and older tended to choose VA hospitals less often than those between 45 and 64 years of age, but more often than those under 45.

In Figure 1 we present those results in the most straightforward way, using odds and odds ratios. Using the fitted model frequencies, we divided the number of non-VA hospital episodes by the number of VA hospital episodes for any particular age, service-connected status, insurance and income-specific category. This produces the odds that a veteran will be hospitalized in a non-VA rather than in a VA hospital. For example, twice as many non-VA hospi-

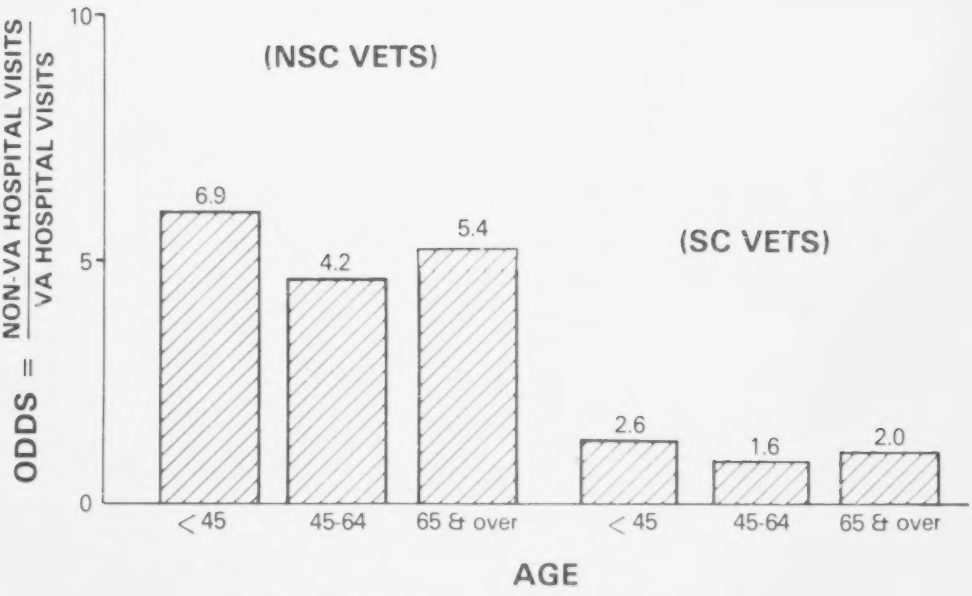


FIG. 1. Odds* of going to non-VA hospital by age.

* Values shown are for low-income veterans with health insurance; however, the same relationships hold over all income and insurance categories.

tal episodes as VA hospital episodes means that the odds are 2 to 1 for non-VA hospitalization. Such odds and even ratios of odds can be expressed mathematically as functions of log-linear model coefficients,[8] but because odds have a simpler substantive interpretation, they should be preferred. Several examples follow.

The following interpretation of Figure 1 leaves out the over-65 age group for simplicity. Reading directly from the graph, it can be seen that service-connected (SC) veterans under the age of 45 are 2.6 times more likely to go to a non-VA than to a VA hospital, and that those 45–64 years of age only 1.6 times as likely. For non-service-connected (NSC) veterans, the odds are 6.9 and 4.2, respectively. Importantly, although these odds differ for service-connected and non-service-connected veterans, the *ratio* of the under 45 years to the 45–64 years odds remains the same for both service-connected and non-service-connected veterans. That is, 2.6 ÷ 1.6 = 1.6 for service-connected veterans, and 6.9 ÷ 4.2 = 1.6 for non-service-connected veterans.

This is an important point. Because the ratios of odds are identical in these two groups, we call it the "age effect," and it is the same regardless of service-connected status. One can also show that this age effect is independent of income and insurance. Thus, we can talk about an age effect without having to specify service-connected status, insurance or income categories, since the odds ratios are the same over these categories.

Health Insurance

In a similar manner we present the data on the health insurance effect. As with age, the effect of health insurance is indepen-

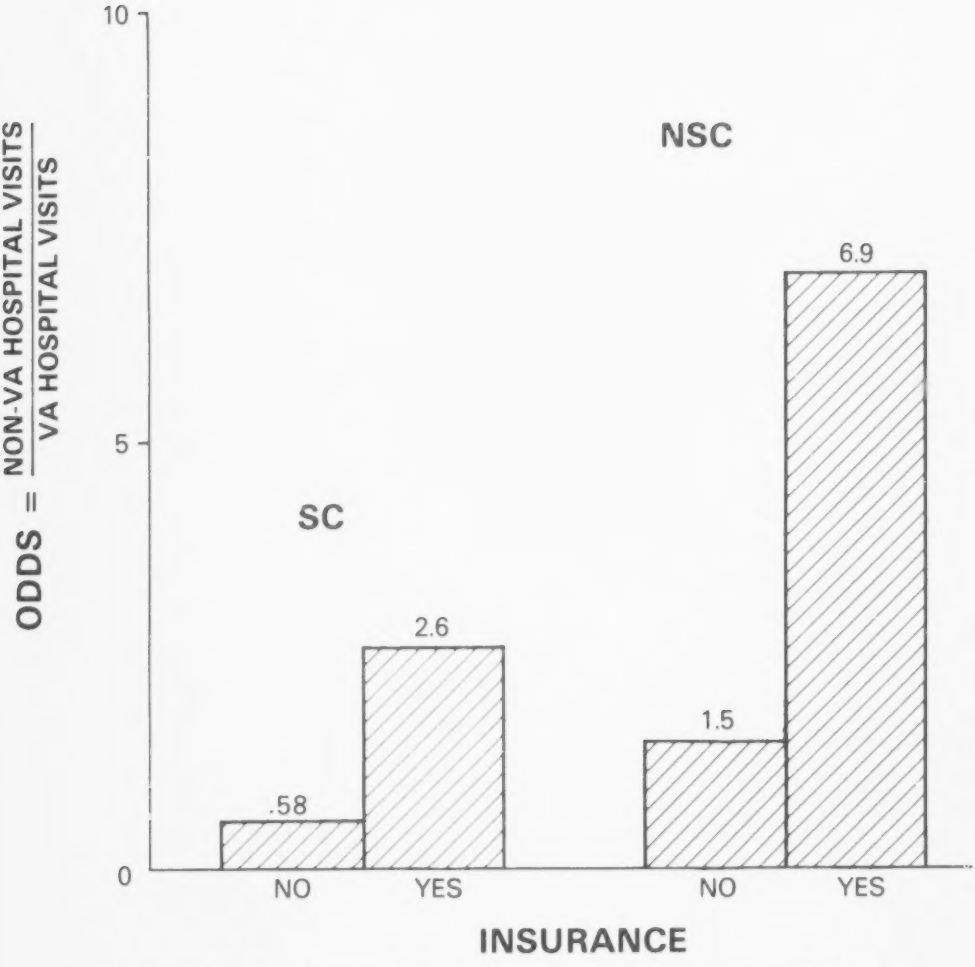


FIG. 2. Odds* of going to non-VA hospital by insurance and service-connected status.

* Values shown are for low-income veterans under age 45 only, however, the same relationships hold over all income and age categories.

dent of all other factors, but we show service-connected status in Figure 2 for comparability. For service-connected veterans with low income and age under 45 years, the odds of going to a non-VA hospital are 0.58 without insurance and 2.6 with insurance; for non-service-connected veterans, they are 1.5 and 6.9 respectively. Once again we note that the odds ratios are the same for service-connected and non-service-connected veterans, namely, $2.6 \div 0.58 = 4.5$ for service-connected veterans, and $6.9 \div 1.5 = 4.5$ for non-service-connected veterans. Hence, we call the common value of the odds ratio (4.5) the "insurance effect."

This odds ratio of 4.5 has an important substantive interpretation. It means that

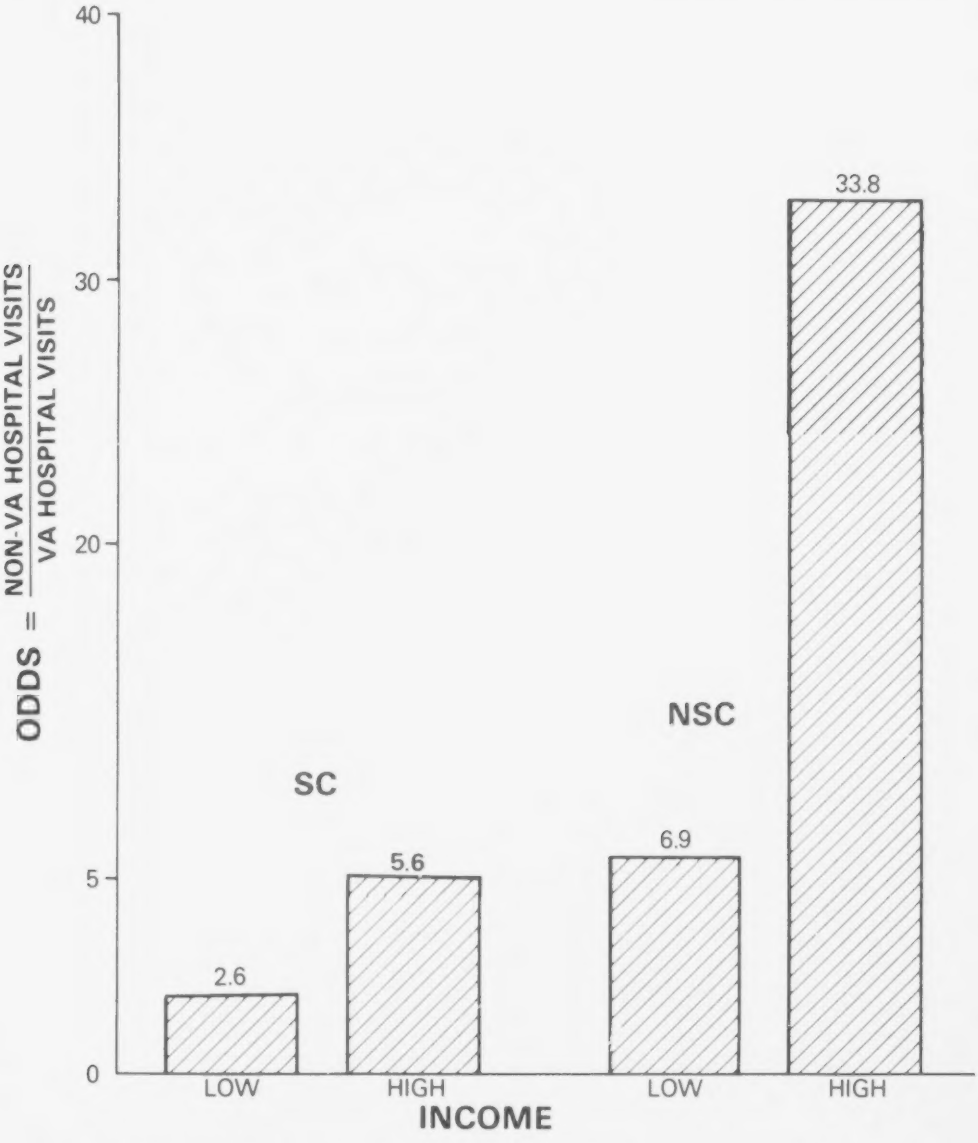


FIG. 3. Odds* of going to non-VA hospital by income and service-connected status.

* Values shown are for insured veterans over age 45, however, the same relationships hold over all age and insurance categories.

veterans with health insurance have 4.5 times higher odds of going to a non-VA hospital than do those without health insurance. (Equivalently, those without health insurance have 4.5 times higher odds of going to a VA hospital.) This is a most important effect on the choice of VA hospitals, and it is independent of age,

service-connected status and income. That is, in any and every age, service-connected status or income category, those without health insurance show the same 4.5 fold odds ratio favoring VA hospitals. The reader should note that this does not mean that VA hospitals are used more than non-VA hospitals in any absolute sense. It means only that no matter what the odds of non-VA hospital usage are in a particular group with health insurance, they are 4.5 times higher in the corresponding group without health insurance.

Income and Service-Connected Status

The effects of income and service-connected status are not as simple and straightforward to present. The reason is that the effect of income varies with service-connected status and the effect of service-connected status varies with income group. Figure 3 illustrates this graphically.

Figure 3 shows how service-connected status and income act together in the choice of a VA hospital. Among service-connected veterans, high-income veterans have more than double the odds of going to non-VA hospitals (5.6 ÷ 2.6 = 2.2). Among non-service-connected veterans, high-income veterans have five times the odds of going to non-VA hospitals (33.8 ÷ 6.9 = 4.9). In this case the odds ratios for service-connected and non-service-connected veterans are *not* equal as they were for both age and health insurance. This indicates that the income effect is dependent upon service-connected status or, in statistical terms, there is an income by service-connected status interaction. One can also say, equivalently, that the effect of service-connection is dependent upon income. Among low-income veterans, non-service-connected veterans have 2.5 times the odds of going to a non-VA hospital (6.9 ÷ 2.6); among high-income veterans, non-service-connected veterans have six times the odds of going to a non-VA hospital (33.8 ÷ 5.6). A discussion of this interaction is presented in the following section.

Table 5 summarizes our odds-ratio interpretation. Table 5 shows that age and income have less strong effects than do insurance, income and service-connected status. On the whole, health insurance is the strongest single factor. Whereas income has a strong effect among the non-service-connected veterans, it has a less strong effect among the service-connected veterans.

Discussion

Before plunging into a discussion of the fitted model, we should first note some limitations of the data, particularly the fact that the data were self-reported. Self-reporting biases did not appear to be serious in most of the cases for which we could check aggregate survey data against aggregate VA figures (privacy restrictions would not allow us to do any individual record matching). For example, according to VA records, there were 942,000 VA hospital discharges in fiscal year 1977. If we subtract the 45,000 deaths, we calculate 897,000 VA hospital discharges versus the Survey of Veterans estimate of 894,000, a remarkable agreement. Similarly, VA hospital records show that 21 per cent of the discharges (one person may account for multiple discharges) were for persons 65 years of age or older; the Survey of Veterans figure was 18 per cent.

Turning to service-connected disabilities, the overall Survey of Veterans estimate of the number of service-connected veterans agreed well with VA records. However, VA records show that 25 per cent of VA hospital episodes involved service-connected veterans versus the Survey of Veterans figure of 38 per cent. Since the overall Survey of Veterans estimate of service-connected veterans agrees well with VA figures, it must be that either (self-reported) VA hospitalized veterans are more apt to report themselves as having

TABLE 5.

Source	Odds Ratio*	Interpretation
1. Age		
a. Under 45 years vs 45–64 years	1.6	Veterans under 45 years old have nearly twice the odds of using non-VA hospitals as those 45–64 years of age.
b. 65 and over vs 45–64 years	1.3	Veterans 65 and over have about equal odds of using non-VA hospitals as those 45–64 years of age.
2. Insurance		
Health insurance vs no health insurance	4.5	Veterans with health insurance have nearly 5 times the odds of using non-VA hospitals as those without health insurance.
3. Income		
a. Service-connected (SC), low income vs high income	2.2	SC veterans with low incomes have more than twice the odds of using VA hospitals as those with high incomes.
b. Non-service-connected (NSC), low income vs high income	4.9	NSC veterans with low incomes have five times the odds of using VA hospitals as those with high incomes.

* The method of model fitting used in this paper does not provide direct estimates of odds ratio reliability; however, all the odds ratios reported are significantly different from 1.0 at the $p < 0.10$ level, so that confidence limits around them exclude the value 1.

a service-connected disability or that (self-reported) service-connected veterans are more apt to remember and report VA hospital episodes. Whatever the case, this over-reporting would tend to exaggerate the VA hospital market share among service-connected veterans.

Turning our attention to the fitted model, perhaps the most striking aspect of the log-linear model analysis is the simplicity of the model. With the exception of service-connected disability status and income, the other factors acted independently. Hence, age and health insurance had independent effects regardless of service-connected status and income.

The service-connected status by income interaction is not difficult to explain. Service-connected veterans are given the highest priority for treatment in VA hospitals, while non-service-connected veterans may enter a VA hospital only if they state under oath that they are unable to defray the cost of hospital charges elsewhere (the "inability to pay" requirement does not apply to veterans 65 years of age and older or those in receipt of VA pension).

Thus, it is easy to see why service-connected status and income effects interact. Both high- and low-income service-connected veterans have high priority for VA hospital care, while only low-income non-service-connected veterans have priority for VA hosptial care. Thus, for non-service-connected veterans we saw that the odds that they will go to a non-VA hospital are much higher for high-income than for low-income non-service-connected veterans simply because of "entitlement." It is a bit puzzling that no age by service-connected status interaction was found, since non-service-connected veterans over 65 years of age are not required to meet the "inability to pay" requirement as explained above. This failure

to detect the service-connected status by age interaction may be due to the small sizes in the 65 years and over age group.**

Conclusion

The findings of this paper may provide useful though limited information to policy makers. Although identification and interpretation of policy-relevant findings are jobs best left to others more expert, the author has been asked by the reviewers at least to speculate on some areas of policy interest. Let us take, as an example, National Health Insurance (NHI).

According to the model, which ignores all real-life factors save age, income, insurance and service-connected status, a national health plan that acts like third-party insurance or Medicare/Medicaid could seriously erode the VA hospital market share. This is because veterans with existing insurance (our definition included Medicare/Medicaid) are far less likely to choose VA hospitalization, regardless of age, income and service-connected status. This speculative hypothesis is partly confirmed by early Medicare experience, which showed a drop-off in VA hospital usage rates in veterans aged 65 years and over just after Medicare went into effect. This original drop-off was only temporary; VA hospital usage rates among older veterans are now back to pre-Medicare levels. Note that, although the model predicts a sharp drop in VA hospital usage rates in the NHI scenario above, the nondynamic model cannot address the question of a subsequent return to previous usage rates.

On the other hand, one can imagine a national plan that would cover veterans by simply eliminating the distinctions in admission priority between veterans, especially between service-connected and non-service-connected veterans. Such a plan would, according to the model, increase the VA hospital market share, especially among high-income veterans, and would have the opposite effect of the first scenario.

In all of this speculation, however, one must remember that the results presented have come from a one-time static model. Without a dynamic component to the model, predicting future events is somewhat adventuresome. Recall in particular that the model predicts a usage rate drop in the first scenario but can say nothing about subsequent usage patterns. Nevertheless, the model does help us to understand the current situation. Others have found that veterans who use VA hospitals are poorer, less likely to have health insurance and more apt to have a service-connected disability.[12] What we have added is a documentation of the relative strengths of these factors and their interrelationships.

References

1. National Center for Health Statistics. The nation's use of health resources 1979. Washington, D.C.: USDHEW, Public Health Service, 1980. (DHEW publication in. 80-1240.)

2. Committee on Veterans' Affairs, United States Senate. A report on the aging veteran. Washington, D.C.: U.S. Government Printing Office, 1978. (95th Congress, 2d Session, Senate Committee Print No. 12.)

3. Reports and Statistics Service, Office of Controller, Veterans Administration. National Survey of Veterans. Washington, D.C.: U.S. Government Printing Office, 1980.

4. Veterans Administration. Federal Benefits for Veterans and Dependents. Washington, D.C.: U.S. Government Printing Office, 1980.

5. Rao CR. Linear Statistical Inference and its Applications. 2nd edition. New York: Wiley & Sons, 1973;417.

6. Halperin M, Blackwelder WC, Verter JI. Estimation of the multivariate logistic risk function: A comparison of the discriminant function and maximum likelihood approaches. J Chron Dis 1971;125:158.

7. Press SJ, Wilson S. Choosing between logistic regression and discriminant analysis. J Am Stat Assoc 1978;699.

8. Page WF. Interpretation of Goodman's log-linear model effects, an odds ratio approach. Soc Meth Res 1977;5:419.

9. Dixon WJ, Brown MB, ed. BMDP-77 Biomedical Computer Programs P Series. Berkeley: University of California Press, 1977.

** A preliminary analysis of the next year of data shows the presence of this interaction.

10. Fienberg SE. The Analysis of Cross-Classified Categorical Data. Cambridge, Mass.: MIT Press, 1977; chapter 6.

11. Mare RD. Social background and school continuation decisions. J Am Stat Assoc 1980;75:295.

12. Committee on Veterans' Affairs, United States Senate. Study of health care for American veterans. A report prepared by the National Academy of Sciences, National Research Council. Washington, D.C.: U.S. Government Printing Office, 1977. (95th Congress, 1st Session, Senate Committee Print No. 4.)

Appendix

TABLE 1A. Fitted Data*

Income	Health Insurance	Age Group	Service-connected	Hospitalized	
				VA	Non-VA
Low	Yes	Under 45	Yes	2.55	6.57
Low	Yes	Under 45	No	8.57	58.96
Low	Yes	45–64	Yes	15.15	23.98
Low	Yes	45–64	No	23.41	99.04
Low	Yes	65 and over	Yes	7.43	14.97
Low	Yes	65 and over	No	18.69	100.73
Low	No	Under 45	Yes	2.27	1.31
Low	No	Under 45	No	19.77	30.42
Low	No	45–64	Yes	28.66	10.15
Low	No	45–64	No	20.16	19.08
Low	No	65 and over	Yes	3.44	1.55
Low	No	65 and over	No	10.41	12.56
High	Yes	Under 45	Yes	3.48	19.35
High	Yes	Under 45	No	5.96	201.51
High	Yes	45–64	Yes	12.37	42.30
High	Yes	45–64	No	15.60	324.26
High	Yes	65 and over	Yes	1.74	7.58
High	Yes	65 and over	No	1.24	32.87
High	No	Under 45	Yes	0.93	1.16
High	No	Under 45	No	2.21	16.72
High	No	45–64	Yes	8.15	6.24
High	No	45–64	No	6.54	30.41
High	No	65 and over	Yes	1.04	1.01
High	No	65 and over	No	0.31	1.86

* Cell values divided by approximate average weight.

TABLE 2A. Raw Data*

Income	Health Insurance	Age Group	Service-connected	Hospitalized	
				VA	Non-VA
Low	Yes	Under 45	Yes	0.85	7.27
Low	Yes	Under 45	No	10.07	56.47
Low	Yes	45–64	Yes	16.91	21.21
Low	Yes	45–64	No	23.75	97.70
Low	Yes	65 and over	Yes	3.66	17.74
Low	Yes	65 and over	No	19.74	98.68
Low	No	Under 45	Yes	1.76	0.82
Low	No	Under 45	No	16.04	33.14
Low	No	45–64	Yes	31.84	5.98
Low	No	45–64	No	17.25	20.99
Low	No	65 and over	Yes	1.48	2.52
Low	No	65 and over	No	11.16	10.82
High	Yes	Under 45	Yes	5.23	16.60
High	Yes	Under 45	No	4.90	201.59
High	Yes	45–64	Yes	7.85	45.82
High	Yes	45–64	No	14.03	324.83
High	Yes	65 and over	Yes	2.43	5.89
High	Yes	65 and over	No	0.77	32.33
High	No	Under 45	Yes	0.64	0.45
High	No	Under 45	No	2.27	15.66
High	No	45–64	Yes	7.51	5.88
High	No	45–64	No	6.89	29.05
High	No	65 and over	Yes	1.06	0
High	No	65 and over	No	0	1.17

* Cell values divided by approximate average weight.

MEDICAL CARE
March 1982, Vol. XX, No. 3

Satisfaction With Maternity Care:
A Matter of Communication and Choice

DEBORAH A. SULLIVAN, PH.D.,* AND RUTH BEEMAN, R.N., C.N.M., M.P.H.†

The purpose of this study was to determine the level of satisfaction with maternity care and whether satisfaction was related to the patterns of communication between caretakers and patient and to specific clinical procedures used during labor and delivery. Data, gathered in a statewide survey, indicated widespread satisfaction with prenatal, labor and delivery care. Level of satisfaction was related to both the quality of communication and the fulfillment of preferences about the management of labor and delivery. There was a general consensus among respondents about the desirability of childbirth coaches, controlled breathing and relaxation techniques, and freedom to move around and the undesirability of medication rendering one unconscious for delivery. There was considerable variation in desire concerning medication during labor, use of a fetal monitor, presence of family members and choice of atmosphere. The implications for alternatives in maternity care are discussed.

JUDGING BY THE FLOOD of articles in the popular women's press, the last five years have been a time of growing unrest with standard maternity care. Numerous authors have questioned such practices as the use of sedatives and analgesics, breaking membranes, inducing labor, episiotomy, fetal monitoring and separation of mother and child.[1–23] All of the "alternatives" suggested involve a greater degree of choice and control on the part of the parents. In particular, the use of childbirth education methods, participant father-coaches, midwives and a home or homelike environment is promoted.

A number of books have been published on these topics. The titles are indicative of the aggressive stand the authors take—*The Rights of the Pregnant Parent*,[24] *Helping Ourselves, Families, and the Human Network*,[25] *Immaculate Deception*,[26] *Birth Without Violence*[27] and *Home Birth*.[28] Whether or not these publications accurately reflect a widespread uneasiness with the interjection of sophisticated technology into the birth process is difficult to establish. Most of the women's literature is based on personal experiences and opinions or limited case studies. Its representativeness is questionable. Only a few consumer surveys of maternity care have been conducted in the United States, and their generalizability also is limited due to sampling techniques or restricted content.[29–32] These same shortcomings also contribute to some discrepancy about desired services, although all report a generally positive evaluation of care.

There are several methodological problems inherent in measuring satisfaction with maternity care. First, patients are reluctant to criticize their caretakers. This problem is intensified with maternity care, during which the arrival of a healthy baby can overshadow negative experiences and

* Assistant Professor of Sociology, Arizona State University, Tempe, Arizona.

† Maternity Care Nursing Consultant, Bureau of Perinatal Health, Arizona Department of Health Services, Tempe, Arizona.

Reprint requests: Deborah A. Sullivan, Ph.D., Sociology Department, Arizona State University, Tempe AZ 85281.

0025-7079/82/0300/0321/$01.00 © J. B. Lippincott Co.

create a favorable environment for subsequent evaluations of care. Nevertheless, a recent evaluation of alternative, indirect measures of satisfaction with medical care failed to show a positive association with other objective indicators of care.[33] Special attention, therefore, must be paid to the level of reported satisfaction when using a purely subjective rating rather than a rating limited to a satisfied–dissatisfied dichotomy. Second, a patient's satisfaction with care is based on his/her perception of several dimensions of caretaker conduct, including technical competence, emotional support and communication.[34] Research has shown that these dimensions are highly correlated and that a patient's perception of technical competence is influenced largely by the other measures of rapport.[35] Third, in spite of the attention given to maternity care by the popular press, many women do not perceive that alternatives could exist in their own course of care. Instead, most are dependent on caretakers to advise them, and recent ethnographic research has shown that only minimal information is given to maternity patients about options available to them.[36] British consumer surveys reveal a similar finding and observe that the desires and objections of a minority deserve special attention in this context.[37] These findings suggest that consumer surveys of maternity care need to be designed to present alternatives objectively and to probe experience with, and evaluation of, specific procedures and the caretaker–patient rapport.

This report is based on a statewide survey of consumer satisfaction with maternity care. The focus is the relationships between expressed level of satisfaction and two aspects of maternity care—communication between the woman and her caretakers and choice of specific clinical procedures. It is hypothesized that the better the rapport and the greater the fulfillment of choice, the higher will be the level of satisfaction

Methods

In the winter of 1979, the Arizona Department of Health Services conducted by mail a survey of all resident women to whom a birth certificate had been issued between October 15 and November 15, 1978. This time period avoided spurious effects in the timing of births due to major holidays and vacation schedules. A January mailing allowed time for the recording of births and the centralized collection of records from outlying areas in the state. Also, it gave the women time to complete postpartum care.

Of the 3,773 questionnaires mailed, 145 were returned undeliverable, and 1,900, or 52.4 per cent of those delivered, were completed and returned by mid-March, 1979. State confidentiality requirements did not allow a follow-up of the nonrespondents. A comparison of the characteristics of respondents with all women in the sample (Table 1) shows an underrepresentation of teenagers and women from low socioeconomic backgrounds. Separate analyses were conducted for each age, educational and ethnic category. The relationships between satisfaction and communication and choices were consistent across categories, although the overall level of satisfaction was lower for the underrepresented groups. The separate analyses are not presented in this report because of repetition and the extensive overlap among the underrepresented categories. The underrepresentation of low socioeconomic groups introduces a conservative bias in the analysis. Expressed level of satisfaction among respondents is higher than that in the population from which they are drawn.

The consumer satisfaction questionnaire included an introductory letter from the Department of Maternal and Child Health, Arizona Department of Health Services. The letter explained the purpose of the study for the Advisory Task Force on

TABLE 1. A Comparison of the Characteristics of Respondents with Women to Whom Birth Certificates Were Issued in Arizona, October 15 to November 15, 1978

	Original Sample		Respondents	
	N	Per Cent	N	Per Cent
Age				
Under 15	7	0.2	4	0.2
15–19	624	16.5	215	11.4
20–24	1,396	37.0	684	36.2
25–29	1,045	27.7	618	32.7
30–34	502	13.3	274	14.5
35–39	171	4.5	87	4.6
Over 40	28	0.7	10	0.5
Unknown	0		8	
Education				
Under 9 years	395	10.5	119	6.3
9–12 years	2,254	59.7	921	48.7
13–16 years	988	26.2	684	36.1
Over 16 years	136	3.6	169	8.9
Unknown	0		7	
Number of children born				
1	1,529	40.5	837	44.1
2–4	1,979	52.4	961	50.6
5 or more	265	7.0	98	5.2
Unknown	0		4	
Ethnic group*				
Anglo	—	72.1	1,418	74.9
Mexican-American	—	18.7	273	14.4
Black	—	2.9	46	2.4
Indian	—	5.5	110	5.8
Other	—	0.6	47	2.5
Unknown	—		6	

* Arizona birth certificates contain information on race but not ethnicity. This does not permit identification of the largest minority—Mexican-American. Therefore the per cent breakdown is based on the ethnic composition of the state in 1975.

Alternatives in Maternity Care and insured confidentiality, since neither the respondents nor their health caretakers could be identified.

The questionnaire, which underwent several pretests and revisions, consisted of 60 structured questions. The questions covered socioeconomic and demographic characteristics of the respondents, prenatal care, labor and delivery care, and postpartum care. Two sets of questions concerned satisfaction with, and quality of, communication with caretakers during the prenatal period and during labor and delivery. Another series of questions asked respondents whether they had specific clinical procedures during labor and delivery and whether or not they had wanted these procedures. In addition, respondents were asked directly for evaluations of their prenatal and labor and delivery care. The rest of the questionnaire concerned postpartum care.[38]

Contingency table analysis was used to investigate the relationship between satisfaction with care and communication patterns and clinical procedures. Associations were evaluated with chi-square tests, and the strength of the associations was measured with Cramer's V. Chi-square values having probabilites of 0.01 or less were considered significant.

Results

Respondents ranged from under 15 years of age to over 40. One quarter were part of a minority ethnic group. Family income varied from less than $5,000 to greater than $20,000. Eighty-eight per cent were married at the time of the survey, and 44 per cent were reporting on their first birth. The vast majority of respondents, 88 per cent, saw a medical doctor for prenatal care, and only 53 of the 1,900 births were delivered outside of a hospital.

Communication during Prenatal Care

The overall evaluation of prenatal care by respondents is extremely positive. Ninety-three per cent reported that they were "satisfied" or "very satisfied" with their care. The majority gave their caretaker favorable evaluations on the three questions about quality of communication during prenatal care (Table 2).

The relationship between quality of communication and overall level of satisfaction with prenatal care can be seen in Table 3. Each of the three measures of communication has a statistically significant relationship with satisfaction with prenatal care. The association is strongest for time spent discussing problems and empathy on the part of the examiner. The less time spent discussing problems and the less empathy shown by the caretaker, the lower is the level of satisfaction.

The association between vocabulary used and level of satisfaction is not as straightforward. The decline in level of satisfaction with the less frequent use of words understood by respondents is reversed for respondents who reported that their caretaker "never" used words they could understand. Perhaps these respondents regarded medical terminology as a reassuring indicator of the expertise of their caretaker. Another possibility is that the medical terminology intimidated women and created a setting in which they were expected to be satisfied regardless of their ability to understand the words used by caretakers.

Communication during Labor and Delivery Care

The findings on labor and delivery care are similar to those on prenatal care. Level of satisfaction is positively skewed (Table 4). An even larger proportion of respond-

TABLE 2. Communication Patterns during Prenatal Care and Labor and Delivery*

	N	Always %	Usually %	Occasionally %	Never %	Do Not Know %	Total %
Prenatal care							
Spent enough time discussing problems	1,887	50	31	14	4	1	100
Used words I could understand	1,880	57	24	10	9	0	100
Tried to understand how I felt	1,877	50	29	11	6	4	100
Labor and delivery							
Explained procedures	1,870	65	24	7	4	0	100
Kept me /support person informed of progress	1,860	68	19	8	4	1	100
Explained choices available	1,843	61	16	6	13	4	100

* Respondents with missing values are not included in the table.

TABLE 3. Satisfaction with Prenatal Care by Perception of Care Received*

	N	Very Satisfied %	Satisfied %	Unsatisfied %	Very Unsatisfied %
Enough time spent discussing problems					
Always	941	81	17	1	1
Usually	576	42	53	4	1
Occasionally	263	18	62	16	4
Never	78	19	28	30	23
Do not know	21	38	43	14	5
Chi square = 191.91				Cramer's V = 0.16	
Used words I could understand					
Always	1,061	70	26	2	2
Usually	449	38	51	9	2
Occasionally	189	35	52	10	2
Never	162	58	33	5	4
Do not know	12	50	25	25	0
Chi square = 62.45				Cramer's V = 0.09	
Tried to understand how I felt					
Always	943	81	18	1	1
Usually	550	45	52	3	0
Occasionally	208	14	70	15	1
Never	99	8	30	38	23
Do not know	71	39	48	11	1
Chi square = 163.11				Cramer's V = 0.15	
Total	1,885	58	35	5	2

* Respondents with missing values or those few reporting mixed satisfaction are not included in the table.

ents felt that their care takers were more responsive to their needs during labor and delivery than during the prenatal period (Table 2). However, it is during this stage of maternity care that individual experiences seem to differ greatly. Among 500 unsolicited comments on the questionnaire were some very positive endorsements, such as "Doctor . . . was fantastic," or "OB staff was excellent." Others wrote of lack of attention, insensitivity and incompetence. For example, one woman wrote that when her water broke the "nurse yelled about getting her wet!" and another wrote, ". . . the doctor was grouchy and unpleasant. He left a sponge in my vagina and didn't stitch me right."

Many of the women offering highly negative comments still checked "satisfactory" in response to the question evaluating overall labor and delivery care. This supports the previously discussed reluctance to criticize caretakers and makes the distinction between "very satisfied" and "satisfied" especially important for this study.

The strong association found between perception of care and level of expressed satisfaction with prenatal care is also evident with labor and delivery care, as seen in Table 4. Respondents who felt that their caretakers did not communicate with them about their labor and delivery experience were far more likely to be unhappy with that experience. Each of the three measures of communication has a statistically significant relationship with satisfaction as measured by a chi-square test.

TABLE 4. Satisfaction with Labor and Delivery Care by Perception of Care Received*

	N	Very Satisfied %	Satisfied %	Unsatisfied %	Very Unsatisfied %
Explained what they were going to do					
Always	1,197	64	30	3	2
Usually	437	43	48	6	3
Occasionally	135	33	45	12	10
Never	77	23	40	16	20
Do not know	8	50	38	12	0
Chi square = 221.78				Cramer's V = 0.17	
Kept me informed of progress					
Always	1,251	64	30	3	2
Usually	347	37	51	8	4
Occasionally	139	31	43	13	13
Never	80	31	38	12	19
Do not know	27	44	48	7	0
Chi square = 229.50				Cramer's V = 0.18	
Explained choices available					
Always	1,116	67	28	3	2
Usually	291	43	48	6	3
Occasionally	113	29	51	11	9
Never	230	31	45	12	11
Do not know	78	44	46	5	5
Chi square = 213.87				Cramer's V = 0.17	
Total	1,876	55	35	5	4

* Respondents with missing values are not included in the table.

Choices during Labor and Delivery

The series of questions about specific procedures during labor and delivery allow us to investigate the extent to which women have preferences concerning the management of events during labor and delivery and whether the fulfillment of such preferences is associated with the level of satisfaction. Respondents who had cesarean births were instructed to answer all of the questions that applied to them before their surgery.

It is clear from Table 5 that women do have preferences about how labor and delivery events are managed. The data show a widespread desire for a childbirth coach and encouragement to use breathing and relaxation techniques, and a desire to avoid medication that renders a woman unconscious for delivery. There is also considerable support for freedom to move around and the use of a fetal monitor. In contrast to the general uniformity of preferences about these procedures, there is a striking divergence of desire about others. These include having family members present, using sedatives and analgesics during labor, and having choice about the atmosphere during delivery.

The extent to which such preferences are honored (no discrepancy between what they wanted or did not want and what they had) varies depending on the procedure (Table 5). When such preferences are not honored, the reported level of satisfaction declines significantly (Table 6). Since few women actually rated their labor and de-

TABLE 5. Percentage Distribution of Respondents According to Procedures Wanted Versus Those Had During Labor and Delivery*

	N	Wanted		Not Wanted		Total Discrepancies
		Had	Not Had	Had	Not Had	
Childbirth coach	1,678	82	9	1	8	10
Encouragement to use breathing/ relaxation	1,621	86	6	2	6	8
Freedom to move around	1,582	62	23	3	12	26
Fetal monitor	1,497	62	12	7	20	19
Family member present	1,648	43	12	2	42	14
Choice of atmosphere	1,379	23	38	2	38	40
Medication to assist me in labor	1,620	50	8	4	38	12
Medication to put me out for delivery	1,668	9	5	3	83	8

* Respondents with missing values for either "wanted" or "had"—because of a cesarean birth, because they did not choose to answer, or because they said they did not know whether they wanted or had a procedure—are excluded from the table.

livery experience negatively, the greater impact of unfulfilled preferences can be seen in the shift toward more women responding that their labor and delivery experience was "satisfactory" instead of "very satisfactory."

Discussion

Contrary to the impressions given by the popular press, our statewide survey reveals widespread satisfaction with maternity care. This finding must be tempered with the acknowledgment that satisfaction levels among our respondents are somewhat higher than those of the population because of the underrepresentation of teenagers and low socioeconomic groups. Also there is evidence that women hesitate to criticize maternity care. Even though some respondents reported that their prenatal caretaker only "occasionally" spent enough time discussing problems, using understandable words or trying to understand how they felt, the majority rated their prenatal care as satisfactory. Similarly, the majority of the relatively few women who said that their caretakers never explained what they were doing to them during labor and delivery, or kept them informed on their progress or explained the choices available to them, still said their care was satisfactory.

One explanation for positive evaluation of such unusually deficient care is the joy surrounding the arrival of a healthy baby. The healthy baby legitimizes the entire pregnancy experience and creates a favorable halo. This undoubtedly results in an upward bias in the expressed level of satisfaction. The halo also may indirectly influence satisfaction by enhancing a woman's evaluation of the quality of communication she had with her caretaker and whether a specific clinical procedure was desirable.

An additional explanation is the more subtle effect of socialization of women as

TABLE 6. Satisfaction with Labor and Delivery Care by Preferences about Specific Procedures*

	N	Very Satisfied %	Satisfied %	Unsatisfied %	Very Unsatisfied %
Childbirth Coach					
Choice honored	1,501	58	34	5	3
Choice not honored	168	41	42	10	6
Chi square = 31.24				Cramer's V = 0.14	
Encouragement to use breathing/relaxation					
Choice honored	1,483	60	33	5	3
Choice not honored	130	35	46	10	9
Chi square = 41.86				Cramer's V = 0.16	
Freedom to move around					
Choice honored	1,169	64	30	4	2
Choice not honored	406	41	44	8	8
Chi square = 81.29				Cramer's V = 0.23	
Fetal monitor					
Choice honored	1,216	60	32	5	3
Choice not honored	274	44	41	8	7
Chi square = 23.92				Cramer's V = 0.13	
Family member present during labor					
Choice honored	1,398	59	33	5	3
Choice not honored	239	41	45	8	6
Chi square = 29.14				Cramer's V = 0.13	
Atmosphere					
Choice honored	830	65	30	4	2
Choice not honored	545	46	40	8	6
Chi square = 58.98				Cramer's V = 0.21	
Medication to assist me to relax in labor					
Choice honored	1,426	61	33	4	2
Choice not honored	188	32	43	12	13
Chi square = 106.27				Cramer's V = 0.26	
Medication to put me out for delivery					
Choice honored	1,537	59	34	4	3
Choice not honored	123	27	41	16	16
Chi square = 11.99				Cramer's V = 0.29	

* Respondents with missing values are not included in the percentages.

potential maternity patients. The rapid declines in maternal mortality during the [illegible] created a social consensus in the [illegible] about the role of medical practitioners as interventionistic caretakers and the role of women as passive recipients of care. This duality brought with it the potential for communication problems and loss of decision making, which most women accept, or at least tolerate, as part of the modern science of obstetrics.

At a time when the popular media are presenting numerous challenges to standard obstetric care, the level of satisfaction with maternity care is an important factor in the future choice of caretaker and type of care. It was concern about the increasing number of home births in the state that led to the present study. Many of the 39 respondents who had a planned home birth attached letters to their questionnaires explaining their choice. Most had previously borne children in a hospital setting. Few related any unusually negative experiences. The most common theme, echoing the popular media, was a desire for a more

personal experience. The women described this desire in terms of greater rapport with their caretaker and more flexibility in procedures, which allowed them more control over their birthing experience. Some of those using the services of midwives in a hospital setting made similar comments.

The results of our study suggest that the desire for rapport and flexibility are not confined to the extremely small number of women resorting to radically different forms of maternity care. Indeed, for the much greater proportion of women who had just completed a more traditional obstetric experience, extent of rapport and fulfillment of clinical preferences were strongly associated with their level of satisfaction. These associations may be inflated to some extent due to the halo effect previously mentioned. Nevertheless, these findings give tentative indication that users of standard obstetric systems of care do not differ greatly from users of radically different systems in terms of the type of relationship they want with their caretakers. They do differ somewhat in their choices of clinical procedures.

Additional analyses, not included in this report, indicate that users of alternative care systems uniformly wanted to avoid medication and other interventionistic practices. In contrast, users of more standard obstetric care systems expressed much greater variation in desire on some of the clinical procedures, especially having family members present, use of medication in labor and choice of atmosphere for childbirth. Some of the variation in expressed desires may be the result of subsequent rationalization of experiences that resulted in a healthy baby.

The data show that the popular women's press does accurately reflect the interests of the majority of women on some clinical procedures and of a sizable minority on other issues. Even allowing for rationalization of experiences, there is much desire for childbirth coaches and controlled breathing and relaxation, and little desire to be rendered unconscious for delivery. A majority still wanted medication during labor in addition to breathing and relaxation techniques, but a large minority, 41 per cent, preferred none. On these four issues, most women's desires were fulfilled. When they were not, level of expressed satisfaction was significantly lower. Of all the preferences, those surrounding uses of medication seem most salient for level of satisfaction.

Many women's choices were not honored on two other alternatives, the freedom to move around and choice of atmosphere (defined on our questionnaire as "dim lights, quiet music, warm bath, pictures, etc."). Taken together with the preference about the presence of family members, the data suggest support for the homelike birthing room alternative advocated in the women's media among more than half the respondents. One hospital in the state that has implemented a birthing room program reports success in averting some home births and enhancing consumer satisfaction.[39]

Implications

Because of the low levels of maternal and infant mortality achieved by modern obstetric practices, it might seem gratuitous to be concerned with consumer satisfaction with maternity care. If satisfaction were based solely on outcome, virtually all consumers would be "very satisfied" with the care they received. The fact that all consumers are not points to the importance of the experience as well as the outcome.

The uneasiness that some consumers have felt with their maternity experience has been transformed into a small but vocal movement seeking to modify current practices. Few reject modern obstetric practices, although they argue for greater choice and control for the parents. Their expressed aim is to personalize maternity care.

The results of this survey suggest that a more personal, more satisfactory course of care depends on more attention to the in-

terpersonal relationships between caretakers and patients and greater flexibility in clinical procedures. To a large extent these two suggestions are intertwined. If caretakers work to establish a good rapport with their patients, they will be advocates for their patients' desires concerning the management of labor and delivery. Similarly, if patients have a good rapport with their caretakers, they will understand the desirability of recommended procedures. Such a multifaceted approach would increase the efficiency of care rather than create more problems. At the same time it would remove some of the motive that has led a small but increasing number of prospective parents to chance less safe procedures.

References

1. Albertson M. Seven choices of delivery. Good Housekeeping 1978;186:68.

2. Ander M. Return of the midwife. Good Housekeeping 1976;183:198.

3. Bing E. Childbirth: liberated labor. Harper's Bazaar 1976;110:103.

4. Brazelton TB. Welcoming your baby: natural childbirth. Redbook 1979;153:35.

5. Clark NH. Traditional alternatives: Lamaze method at maternity wing of New York Hospital Cornell Medical Center. Harper's Bazaar 1976;110:103.

6. Davis F. Who's having the baby—you or the doctor? Woman's Day 1976;March:10.

7. Frank A, Frank S. Home or hospital delivery. Mademoiselle 1978;84:66.

8. Gardner J. Having your baby at home. Glamour 1979;77:240.

9. Hoover MB. Hidden rewards of childbirth training. McCall's 1974;101:59.

10. Lake A. Childbirth in America. McCall's 1976;103:83.

11. Marshall MC. My husband delivered our baby. Good Housekeeping 1976;179:77.

12. Maynard F. Home births vs. hospital births. Women's Day 1977;June:10.

13. Milinaire C, LeBoyer F. Born happy. Vogue 1974;164:80.

14. Norwood C. Humanizing way to have a baby: with nurse midwives at birth centers. MS 1978;6:89.

15. Panter GG. Prepared childbirth—moving ahead. Parents 1979;54:33.

16. Pascoe ES. How women want to have their babies. McCall's 1977;105:109.

17. Powers T. Fathers in the delivery room. Good Housekeeping 1974;179:93.

18. Randal J. Where should your child be born? Parents 1979;54:74.

19. Rose L. Pregnant women's guide to pain relief for labor and delivery. Glamour 1980;78:152.

20. Samuels M, Samuels N. Relax and visualize your way to a better pregnancy and birth. Parents 1979;54:48.

21. Torres T. From womb to world: a nonviolent transition: the LeBoyer method. MS 1974;3:22.

22. Trezise J. Midwives move into mainstream. Working Woman 1978;3:63.

23. Yuncker B. Delivery procedures that endanger a baby's life. Good Housekeeping 1975;180:56.

24. Elkins V. The rights of the pregnant patient. New York: Two Continents Press, 1976.

25. Howell M. Helping ourselves, families and the human network. Boston: Beacon Press, 1975.

26. Arms S. Immaculate deception. Boston: Houghton-Mifflin, 1975.

27. LeBoyer F. Birth without violence. New York: Random House, 1976.

28. Petty R, ed. Home birth: a complete, authorative guide on the back to basics trend in childbirth. Chicago: Domus Books, 1979.

29. Nunnally D, Agular M. Patients' evaluation of their prenatal and delivery care. Nurs Res 1974;23:469.

30. Scaer R, Korte D. MOM survey: maternity options for mothers—what do women want in maternity care. Birth Fam J 1978;5:20.

31. Anderson SF, Kay MA, Shaw E.. Childbirth as they like it. Arizona Nurse 1978;May/June:4.

32. Norr KL, Block CR, Charles A. Explaining pain and enjoyment in childbirth. J Health Soc Behav 1977;18:260.

33. Steward M, Wanklin J. Direct and indirect measures of patient satisfaction with physicians' services. J Community Health 1978;3:195.

34. Ware JE, Doyle BS. Physician conduct and other factors that affect consumer satisfaction with medical care. Publication No. D-5670. Santa Monica: Rand Corporation, 1977.

35. DiMatteo MR, Hays R. The significance of patients' perceptions of physician conduct. J Community Health 1980;6:18.

36. Riley E. What do women want? The question of choice on the conduct of labor. In: Chard T, Richards M. Benefits and hazards of the new obstetrics. Philadelphia: JB Lippincott Co, 1977.

37. Danziger SK. Treatment of women in childbirth: implications for family beginnings. Am J Public Health 1979;69:900.

38. Sullivan DA, Beeman R. Satisfaction with post-partum care: opportunities for bonding, reconstructing the birth and instruction. Birth Fam J 1981;8:153.

39. Kieffer MJ. The birthing room concept at Phoenix Memorial Hospital: consumer satisfaction during one year. JOGN Nurs 1980;9:151.

MEDICAL CARE
March 1982, Vol. XX, No. 3

Utilization of Specialized Ambulatory Care by the Elderly

A Study of a Clinic

STANLEY R. INGMAN, PH.D.,* AND IAN R. LAWSON, M.D.†

Specialized ambulatory geriatric care remains an understudied topic. This article analyzes the organization of one special clinic that was established in an apartment block of elderly housing. The findings indicate that as the number of nurse practitioners and case workers' visits rose steadily during the first four years of the clinic's existence, visits to the clinic physicians declined by approximately one visit per client per year. The clinic population utilized less inpatient hospital care than other nonusing residents, and less than other residents of similar age in Connecticut. There was reduced usage of intermediate nursing home care (as opposed to skilled nursing care). Hence, how geriatric ambulatory care is delivered affects institutionalization. In the interests of general well-being and economy, more attention should be paid to this extramural phase of medical care of the elderly.

CRITICS OF THE HEALTH CARE SYSTEM for older adults have charged that it is biased toward institutionalization, whether it be in the hospital or in the nursing home.[1,2] With the degree of sickness and disability prevalent among the elderly, alleviation of disability and prevention of its more severe, dependency-inducing forms may be expected to have effects on institutional admission rates. There has been marked interest in home care services and improved socioclinical assessment as part of community, state and federal strategies to decrease institutionalization.[3] Ambulatory geriatric care has not received the same attention, yet may have similar potential for reducing the utilization of institutional services.[4,5]

Medicare has concentrated federal and private insurance reimbursement on acute hospital care, as Medicaid has supported the expansion of nursing home care. National data indicate that those aged 65 years and over now occupy slightly less than 40 per cent of the hospital beds on any one day. From 1967 to 1975 there has been a 61 per cent increase in hospital admissions for those aged 65 years and over, as opposed to only 16 per cent for those under 65 years of age.[6]

This increase in the utilization of hospital services by the elderly is probably the result of two major processes: 1) the increased availability of these services to the elderly because of Medicare and 2) their financial advantage to the using physician.

From the Department of Family and Community Medicine, University of Missouri-Columbia, School of Medicine,* and the Department of Community Medicine and Health Care, University of Connecticut, Storrs, Connecticut.†

Supported in part by the Community Studies Unit, Department of Community Medicine and Health Care, UCONN-Health Center.

Dedicated to Sister Barbara Freeman, Nurse Practitioner in the program, who died September 9, 1980. The authors appreciate the careful comments and suggestions from Derek Gill, Ph.D., and Mike Stotts, M.S. Leonard Menczer, D.D.S., M.P.H., Hartford City Health Department, was instrumental in the program's initiation and growth.

Reprint requests: Stanley R. Ingman M-330A, University of Missouri-Columbia, School of Medicine, Columbia, MO 65201.

0025-7079/82/0300/0331/$00.90 © J. B. Lippincott Co.

Studies of Ambulatory Care for the Aged

Wan in 1977 reported a study of 2,000 noninstitutionalized older persons living in low-income urban areas near five different neighborhood health centers. Those patients in the catchment areas surrounding the neighborhood health centers increased their utilization of ambulatory medical care by a factor of three as compared with a similar group of elderly persons outside the area covered by these facilities.[7]

In Baltimore, German et al. point out that the type and nature of ambulatory care delivery does increase the likelihood that elderly individuals will request care for less life-threatening but potentially serious conditions. In addition, the elderly using the health maintenance organizations in Baltimore are more likely to be seen for symptoms than were those receiving care through the traditional system of medical care, e.g., physician office practice or hospital outpatient departments.[8]

Davis and Reynolds (1976), using the National Health Survey data, showed that health status is the most important predictor of physician and hospital utilization and that those that are ill use more services. This challenges the commonly held opinion that if services are made "too" accessible, the "well-lonely" elderly will overutilize those services.[9]

Barney and Neukom argue that the elderly need not only more services, but that they need specialized health services.[10,11] In Michigan they observed that many elderly seem to be rather circumspect and pessimistic about conventional health care, i.e., they commonly hold the view that "efforts to ameliorate the symptoms of decline involve excessive cost for questionable benefits.[11] Such health beliefs (adopted by many health care providers as well as elderly) according to Barney and Neukom (1979), serve to explain the national problem of meager provisions and meager use of geriatric care. Barney and Newkom describe the main points of that belief system: "old age is a malady in which all are susceptible and all will succumb, in treating its symptoms health care is increasingly costly and increasingly ineffective."[10] Several actions to combat these beliefs and improve the provisions of geriatric services are outlined by Barney and Neukom.[11]

> First, because it is not clearly known that health services are of true benefit to the older patients, more programs like the Pilot Geriatric Arthritis Program are needed to carry on demonstrations and evaluations. Second, on the basis of what is known, policy planners should allocate resources to health services for the elderly, without fear that the patient demand will get out of control. Finally, policy planners should consistently allocate a portion of the geriatric health care resources, as did the Arthritis Program, to designing, publicizing, and delivering services, in such a way as to reach the isolated, the needy, and the reluctant elderly patient.

By these means, the authors feel that geriatric health care could play a vital role in assuring health and independence to elderly Americans.

Development of specialized geriatric health care services may be supported from the perspective of continuity of service provision. While continuity of care is considered to be one of the major advantages to be derived from the provision of specialist primary patient care, continuity of care is even more apposite for the care of the elderly. Here, the mix of social and medical factors in illness are complex, and continuity in the provider system is essential for both understanding and management. In this article we examine a population of elderly who receive ambulatory care service where there were no out-of-pocket costs in terms of providing health maintenance, continuity of care, a range of social and clinical support services, and "routine" medical care. The description concentrates on the quality and type of medical care utilized by the participants in the Smith Tower Clinic.

The Geriatric Medical Clinic

In 1972 at the invitation of the Hartford City Health Department, the clinic was established in a 200-apartment building for low-income elderly. The Housing Authority was permitted to allocate two apartments to this clinic and to a podiatry/dental clinic. A secretary, a nurse practitioner and a case coordinator were the full-time staff five days a week. Typically, physicians' sessions were held three mornings a week. Medical and social histories based on a protocol were taken by the nurse practitioner, either in the patients' homes or on their first visit to the clinic. Physical examination was also conducted by protocol.

Elderly residents in the apartment building and in the community were told that attendance at the clinic required a commitment to the clinic program, that is, they were to view the clinic as the sole service for primary care and specialty referral.

After a comprehensive history and examination were completed by the nurse practitioner, every patient was seen by an involved physician who made selective follow-up examinations. Each visit to the nurse practitioner case coordinator and physician was recorded in the clinical record, as were admissions to other facilities in Hartford. The record format was problem oriented, with taxonomic ordering[12] of the master problem list; drugs being taken, target symptoms and selected laboratory results were listed on a flow sheet. Up-to-date copies of these were available at the related hospital's emergency rooms and accompanied patients when they were referred out to specialty offices and clinics.‡

After three months as a free-standing clinic, the clinic was incorporated into the Mt. Sinai Hospital outpatient department. This allowed the clinic to receive reimbursement from Medicare and Medicaid and to ensure better integration with hospital emergency services. The city continues to contribute funds to permit the hospital to waive Medicare co-insurance and deductible charges for all elderly patients. Patients therefore meet only the costs of medication. Private specialty consultants agreed to accept assignment and frequently waived billing for co-insurance and deductible. The Hartford Foundation for Public Giving provided basic office equipment.

Clinical records form the basis of the analysis presented in this article. In 1976 the records from 1972–1975 were coded and processed for computer analysis. The records included data derived from the first visits, together with information presented at subsequent visits. More research funding was secured in 1977, so a similar process of coding was done for patients entering the clinic in 1976 and 1977. Therefore, when analyzing certain outcome factors such as presem status, we will separate the two clinic populations in our analysis (i.e., 1972–1975 group and 1976–1977 group).

A prospective comparative study of the aged in various high-rise apartments was being done by Dr. Adrian Ostfeldt of the Yale School of Public Health during the 1970s. Since some of his subjects became patients at our clinic, we were able to compare how their health fared in relation to that of others not served by our clinic program.

Characteristics of Clinic Users

Between May 1972 and December 1977, approximately 644 patients attended the clinic. Two thirds of the clients were women; 49 per cent of the clients were between 71 and 80 years of age; 21 per cent were 81 years old or older; and 30 per cent were under the age of 71 years. Although the clinic was advertised as a geriatric clinic, 13 people under the age of 60 years

‡ Full operational description of the clinic and the nurse practitioner role is available from the authors.

were attracted or referred to the clinic. Approximately 17 per cent of those entering the clinic between 1972 and 1977 were black elderly. With respect to living arrangements, 65 per cent lived alone upon entry into the clinic, and 25 per cent were married. Between 1972 and 1977, 82 per cent of the clients seen reported a monthly income of $300 or less. Slightly less than 10 per cent of the clients reported receiving $600 or more per month.

In the initial years of operation, more than 80 per cent of the clients lived at the Smith Tower apartment building (73 per cent) or within four blocks of the Towers (7 per cent). By 1975, more than 80 per cent of the new patients entering the program lived outside the four-block area. In 1976 and 1977, of the 150 new patients entering the program, only 27 came from within the immediate four-block area. By December 1977, five years after the introduction of the clinic, of the 250 residents in the Smith Towers, 153 had attended the clinic; 96 of this group began attending in 1972.

Between 1972 and 1977 the proportion of those newly registering (a low of 20 per cent in 1973 and a high of 72 per cent in 1977) stated that they previously had a private physician. Nevertheless, throughout the six-year period, only 50 per cent of the new registrants reported seeing a physician in the preceding six months, and 12 per cent stated that they had not seen a physician for more than three years. It would seem that as the clinic became better known to the tenants of Smith Towers and the surrounding area, patients were attracted to the clinic even when they already had a private physician.

Thirty-six per cent had seen a professional about glasses in the year prior to their joining the clinic. Most had not visited a podiatrist (65 per cent), and only 18 per cent had been to a podiatrist in the previous year. Only 25 per cent had been to a dentist in the year prior to their joining the clinic.

Health Status of Clinic Users

Population at Intake. The incidence of selected disabilities of daily living at intake is presented in Table 1. These disabilities, the incidence of particular symptoms, and use of hypnotics and stimulants (see Table 2) are derived from direct leading questions per protocol.

This high incidence of symptoms is commensurate with half of the new registrants in each year declaring themselves to be in only "fair" or "poor" health.[13,14]

To summarize, the clinic population consisted of people living alone with low incomes; as a group, while ambulatory, they carried a high "load of morbidity." In time the clinic attracted more people living farther from the clinic each year and more people who had a regular private physician prior to attending the clinic.

Utilization of Clinic Services

Over four years (1972–1975) of the clinic's operation, as shown in Table 3, there was a decline in visits to the physicians—from an average of 4.0 visits per patient in 1972 to an average of 3.3 visits per patients in 1975.§ In contrast, visits to the nurse practitioners increased every year. In 1972, there was an average of 2.9 visits per year to the nurse practitioner, and an average of 4.8 visits per year in 1975. Twenty-seven per cent of patients in 1972 and 60 per cent in 1975 visited the case coordinator. Those visiting the case coordinator at least once averaged 1.3 visits from 1972 to 1974. In 1975, this rate rose to 5.3 visits per patient to the case coordinator.

Thus, in the initial four years of the clinic's operation there was a decline in physician utilization. In addition, from 20 per cent to 26 per cent of the active patients

§ The National Center for Health Statistics reported that there were on average 4.9 visits to a physician's office by persons 65 years and older (NIA, 1976).

TABLE 1. Levels of Disability Upon Entry (N = 644)

Activities of Daily Living	1972-1977
Difficulty walking up or down stairs	37%
Difficulty cutting toe nails	35%
Difficulty walking	31%
Difficulty washing/bathing	13%
Help to go to toilet needed	4%

(1972–1975) each year did not choose to visit or were not asked to visit one of the clinic physicians. For those patients who see the nurse practitioner and case coordinator, the average number of visits increased steadily until 1976.

In the 1976–1977 cohort, the utilization rates for the 1976 intake (N = 82) were 2.9 physician visits, 3.1 nurse practitioner visits and 3.5 case coordinator visits per year.

In 1977, the average number of visits for patients who entered in 1976 as well as in 1977 was 3.5 visits to the physician at the clinic. In contrast to 1973, there were fewer visits on the average to the nurse practitioner in 1977 ($\bar{X}$ = 2.4). Case coordinator visits also dropped off to an average of 1.2 per patient (range = 1–20) in 1977, as did visits to the physicians at the clinic in 1977. One possible explanation for the decline in clinic utilization relates to pressure from third-party intermediaries for the clinic *not* to see residents for health maintenance or chronic disease follow-up visits.

TABLE 2. Single Symptoms and Selected Medications Taken at Time of Assessment and Entering Clinic (N = 644)

Single Symptom/Drug Taken	1972–1977
Tired	58%
Poor sight	55%
Dental problems	46%
Shortness of breath	44%
Poor sleep	42%
Taking sleeping pills	19%
Use of alcohol	40%
Use of tobacco	39%
Constipation	39%
Taking laxatives	36%
Deafness	35%
Hearing aid worn	5%
Leg swelling	31%
Falls	28%
Unsteadiness on feet	28%
Skin problems	27%
Weight loss	25%
Poor appetite	23%

Many factors influence the increase or decrease of visits to specific occupational groups in the clinic. For example, clinic and insurance company policy, as well as patient behavior, are only three factors. However, the crucial issue is to assess the impact of different visitation patterns on the utilization of other services and on general well-being.

Utilization of Institutional Services by Clinic Participants

Based on the recollection of the patients, the average percentage hospitalized for the year prior to joining the Smith Tower Clinic was 17 per cent from 1972 to 1977 (see Table 4). From clinic records in 1972, it was determined that 47 per cent of those clients who were active in the clinic in 1972 were admitted to a community hospital in Hartford. Of those patients who entered in 1972 and who were still clinic attenders in 1975, 22 per cent were admitted to a hospital. Higher levels of morbidity of early cohorts may explain the initially high hospitalization rates of 47 per cent and 30 per cent. Even with increased access to ambulatory care at the Smith Tower clinic, the percentage of patients hospitalized did not rise, but declined in time. The percentage of patients hospitalized in Connecticut for all persons aged 65 and older was 26 per cent in 1976.[15] In general, the clinic population seemed to utilize hospital services significantly less, both before and after joining the clinic program, than the typical older adult in Connecticut.

On the issue of the average length of stay in the hospital, statistics fluctuated from year to year. The average length of stay per

TABLE 3. Average Number of Visits to Smith Tower Clinic: Physicians, Nurse Practitioners and Case Workers*

Number of Active Patients by Year	Number of Visits to Physicians	Number of Visits to Nurse Practitioners	Number of Visits to Case Workers
1972–1975 Cohorts			
N = 133/1972	4.0	2.9	1.7
N = 231/1973	4.5	4.0	.9
N = 369/1974	3.5	4.6	1.5
N = 410/1975	3.3	4.8	5.3
1976–1977 Cohorts			
N = 82/1976	2.9	3.1	2.5
N = 129/1977	3.5	2.4	1.2

* Calculation includes residents who may not have visited a physician, nurse practitioner or case worker in a specific year.

admission per year ranged from 12.2 to 16.9 days (1972–1975). For the 1976–1977 cohorts, the average length of stay for active patients was 13.0 in 1976 and 11.0 in 1977. The average length of stay in Connecticut during this time was 11.9 days for persons admitted who were aged 65 or older (CHA Reports, 1978). This may suggest that clinic patients who did require admission to a hospital were more ill than the typical Connecticut older adult admitted to a hospital, and thus spent more time when admitted.

TABLE 4. Hospitalization of Clinic Patients

Number of Active Patients Per Year	Percentage Hospitalization	Average Length of Stay
1972 (N = 133)	47	12.2
1973 (N = 231)	30	13.8
1974 (N = 369)	22	16.6
1975 (N = 410)	22	16.1
1976–1977 Cohorts		
1976 (N = 82)	20	13
1977 (N = 129)	18	11

With respect to nursing home use, on the average approximately 2 per cent were admitted each year (1972–1977). In the 1976 cohort, two patients were admitted out of 82 patients (2.4 per cent). In 1977, only three patients were admitted to a long-term care facility out of the 134 active patients who entered the clinic in 1976 and 1977. It is important to note that all nursing home admissions were made to skilled nursing facilities (SNF), and not to facilities providing intermediate-level care. By fully utilizing medical and social services, patients were in considerable dependency before it was thought that long-term institutionalization was warranted. In other circumstances they might have required earlier institutionalization, but to a lower (intermediate) care level.

In 1972, as previously noted, Dr. Adrian Ostfeld was doing a study of the effects of relocation of the elderly. He collected four samples of 50 elderly each in Connecticut. Of the 50 cases collected at the Smith Tower, 30 were nonattenders at the clinic and 20 were attenders. Two years later, in 1974, he noted the following: In the 30 nonattenders, 14 had been hospitalized at least once, two were admitted to Skilled Nursing Care, and there were six deaths. However, in the 20 attenders, only five had been hospitalized, two had been admitted

to Skilled Nursing Care and, most striking, there were no deaths in this group. In reviewing the health status of each group using standard measures, he found no striking differences between the two groups in 1972.

Status of Patients as of June 1976

Approximately 53 per cent of those who entered the clinic between 1972 and 1975 were still active patients as of June 1976. Table 5 indicates that approximately 6 per cent died and that 5 per cent were residing in an SNF as of June 1976. Eighteen per cent chose to seek another source of primary care. The status of 10 per cent was unknown as of June 1976.

The most significant cohort for review as of June 1976 comprised those patients who had entered the clinic in 1972. Of 131 patients, as of June 1976, 51 per cent (67) were still active in the clinic; 12 per cent had chosen to go elsewhere for primary care; 16 per cent had moved out of Hartford or their status was unknown; 9 per cent had died during the four-year period; and 10 per cent were in skilled nursing facilities.

Conclusion

As Wan[7] and German et al.[8] also documented, better outpatient service accessibility tends to increase visitations. The high levels of morbidity and disability among the Smith Tower clinic population should help dispel the notion that the "worried well" will dominate geriatric ambulatory care when services are made freely accessible.

With patients seeing the nurse practitioner, the case coordinator and the physician on different visits to the clinic, the average number of visits to the Smith Tower Clinic was well above the average of 4.9 physician visits per older adult in New England (1973–1974). At the Smith Tower Clinic there was a steady decline in physician visits and an increase in the less costly nurse practitioner and case coordinator visits. As Quinn[16] suggested, when comprehensive geriatric assessment and socioclinical support services are provided to the elderly, less physician services and hospital care are required by the older adult.

Our study investigated the relationship between aggressive, comprehensive ambulatory geriatric care and hospital utilization. Our review of hospitalization prior to entry into the clinic shows that the clients had a lower average rate of hospitalization than that found within the state of Connecticut in 1976. Since poor people are more likely to avoid medical services when copayment is required, this low rate of hospitalization may not be surprising given the large number of poor elderly in our clinic population. At least this may explain the low rates prior to their joining the

TABLE 5. Status of Patients as of June 1976 by Cohorts (1972–1975)*

	1972	1973	1974	1975	Total
Active	67 (51.1%)	68 (61.8%)	90 (58.9%)	54 (53%)	279 (53%)
Moved	7 (5.3%)	6 (5.4%)	5 (3.3%)	6 (6.7%)	24 (5.5%)
Died	12 (9.2%)	5 (4.5%)	5 (3.3%)	3 (3.4%)	25 (5.7%)
SNF	13 (9.9%)	5 (4.5%)	2 (1.3%)	3 (3.4%)	23 (5.3%)
Changed primary care source	16 (12.2%)	20 (18.2%)	33 (21.6%)	12 (13.5%)	81 (18.5%)
Referred to other physician	2 (1.5%)	1 (.9%)	2 (1.3%)	2 (2.3%)	7 (1.6%)
Unknown	14 (10.7%)	5 (4.5%)	16 (10.5%)	9 (10.1%)	44 (10.1%)

The 1976 and 1977 cohorts were not included for this analysis.

clinic. However, clinic membership did increase ambulatory visits when copayments were covered by grant funds and did not cause a rise in hospitalization, as national trends would have suggested. If one notes that hospitalization for the age group nationally was increasing by approximately 5 per cent per year from 1972 to 1975, it is even more significant that the rate for the clinic population did not rise each year.[6] It is possible that intensive ambulatory care and coordination of other services secured for the attenders were able to prevent some hospitalization. Our higher than average "length of stay" for patients hospitalized, and the use of skilled nursing beds as opposed to intermediate care beds, suggest that institutional care was reserved for more dependent or ill patients.

As various community models that focus on comprehensive case assessment and management are being systematically developed and evaluated nationally, the type of geriatric ambulatory care clinic that we described in this article should also be carefully investigated. Current federal policy to establish special services for the frail elderly and to prevent inappropriate institutional admissions should deal with both outpatient care and home care. We would argue that aggressive and more broadly conceived ambulatory care systems can respond to some of the basic needs of the frail and not so frail older adults. Health care finance administrators, at a minimum, should review the Medicare administrative policy of discouraging preventive outpatient care as has been experienced by the staff when third-party intermediaries visited the Smith Tower Clinic.

References

1. Somers AR. The geriatric imperative: a major challenge to the health professions. In: Somers AR, Fabian DR, eds. The Geriatric Imperative. New York: Appleton-Century-Crofts, 1981.

2. Brody SJ. Comprehensive health care for the elderly: an analysis. Gerontologist 1978;13:412.

3. The T.G. interview: Robert Benedict. Gerontologist 1980;20:131.

4. Somers AR. The high cost of health care for the elderly: diagnosis, prognosis, and some suggestions for therapy. J Health Polit Policy Law 1978;3:163.

5. Bellin SS, et al. Impact of ambulatory-health-care services on the demand for hospital beds. N Engl J Med 1969;28:808.

6. Utilization data for senior citizens. Hospitals 1976;50:47.

7. Wan T. Access to care, health status and health service utilization of non-institutionalized older persons in low-income urban areas. Paper presented at Gerontological Society meeting, San Francisco, November 18–22, 1977.

8. German PS, et al. Health care of the elderly in medically disadvantaged populations. Gerontologist 1978;18:547.

9. Davis K, Reynolds R. The impact of medicare and medicaid on access to health care. In: Rosett RN, ed. The role of health insurance in the health service sector. New York: National Bureau of Economic Research, 1976.

10. Barney JL, Neukom JE. Elderly users of health services. Paper presented at Gerontological Society meeting, San Francisco, November 18–22, 1977.

11. Barney JL, Neukom JE. Use of arthritis care by the elderly. Gerontologist 1979;19:548.

12. Lawson IR. A taxonomy of geriatric care using the Problem-Oriented Record System. In: Lawson I, Ingman S, eds. The language of geriatric care: implication for professional review. New Haven: Connecticut Health Services Research Series, 1975.

13. Maddox GL, Douglas EB. Self assessment of health: a longitudinal study of elderly subjects. J Health Soc Behav 1973;14:87.

14. Linn BS, et al. The very old patient in ambulatory care. Med Care 1978;609.

15. Impacts of an aging population on utilization and bed needs of Connecticut hospitals. CHA Reports 1978;1:1.

16. Quinn JL. Triage-coordinated care for the elderly. J Continuing Education in Nursing 1979;10:7.